TECHNIQUES AND ERRORS IN MEASURING CROSS-CORRELATION AND CROSS-SPECTRAL DENSITY FUNCTIONS

R. D. Kelly
L. D. Enochson
L. A. Rondinelli*

*Consultant to Contractor

MAC 511-01
February 1966

MEASUREMENT ANALYSIS CORPORATION
10960 Santa Monica Boulevard
Los Angeles, California 90025

FOREWORD

This report was prepared by the Measurement Analysis Corporation, Los Angeles, California, for the Unsteady Aerodynamics Branch, Aero-Astrodynamics Laboratory, George C. Marshall Space Flight Center, National Aeronautics and Space Administration, Huntsville, Alabama, under Contract NAS 8-20179. Mr. L. Schutzenhofer was the project engineer.

The authors gratefully acknowledge their appreciation for the assistance given by G. P. Thrall and P. H. White of the Measurement Analysis Corporation.

Dr. L. A. Rondinelli, an independent consultant, conducted the survey of the radar and communication fields described in Section 6.

ABSTRACT

23595

This report discusses techniques for the measurement of cross-correlation and cross-spectral density functions and errors in these measurements. Analog, digital and hybrid data reduction techniques presently in use are described. In addition, those techniques used in the radar and communication fields that appear useful for the specific application of interest (dynamic pressure measurements) are detailed.

The error formulae cover both statistical uncertainty errors from the analysis of records with finite bandwidth and finite duration, and hardware related errors. Particular emphasis is placed on phase errors throughout the entire measuring system from the transducer through the analyzer. Both static and dynamic phase errors are covered. In addition, errors from analyzer parameters (filter bandwidth, scan rate, etc.), extraneous additive noise, finite transducer size and magnetic tape recorder velocity nonuniformities are discussed.

Author

TECHNIQUES AND ERRORS IN MEASURING CROSS-CORRELATION AND CROSS-SPECTRAL DENSITY FUNCTIONS

CONTENTS

Foreword ii

Abstract iii

1. Introduction 1

2. Definitions of Correlation and Spectral Density Functions 3

3. Analog Data Reduction Techniques 8

3.1 Correlators 8

3.1.1 Time Delay Mechanisms 10

3.1.2 Multiplication Mechanisms 13

3.1.3 Integration Mechanisms 17

3.2 Cross-Spectral Density Analyzers 21

3.2.1 90^{o} Phase Shift Mechanisms 24

3.2.2 Filter Mechanisms 27

3.3 Conclusions and Recommendations 32

4. Digital Data Reduction Techniques 34

4.1 Introduction 34

4.2 Correlation Function Computations 36

4.2.1 Basic Correlation and Spectrum Computational Method 37

4.2.2 Factoring of Common Terms 41

4.2.3 Eight Level Quantization Method 43

4.2.4 One-Bit Quantization Methods (Extreme Clipping Methods) 45

4.2.5 Sum of Squares Method 48

4.2.6 Direct Computation of Spectrum from Fourier Transform 50

4.2.7 Cooley-Tukey Variation of the Direct Fourier Transform Method 52

CONTENTS (continued)

4.3 Direct Spectral Computations 58
4.3.1 Basic Filtering Procedure 58
4.3.2 Use of the Sin x/x Filter 67
4.3.3 Single Tuned Filter Methods 70
4.3.4 Variable Bandwidth Procedures 73
4.4 A Special Digital Statistical Correlation System 78
4.5 Conclusions and Recommendations 79

5. Hybrid Data Reduction Techniques 86
5.1 General Purpose Hybrid Computer 86
5.2 Special Purpose Hybrid Computers 87
5.2.1 Time Compression Techniques 87
5.2.2 Delta Modulation Techniques 90
5.2.3 Analog Sampling Technique 94
5.3 Conclusions and Recommendations 96

6. Applicable Radar and Communication System Techniques 97
6.1 Introduction 97
6.2 System Requirements 97
6.3 Reduction of Data Requirements 102
6.4 High Capacity Storage Techniques 105
6.4.1 Magneto-Optical Recording 105
6.4.2 Cathode Ray Storage Tubes 107
6.4.3 High-Speed High-Capacity Magnetic Disc Recording 110
6.5 Cross-Correlation Techniques 111
6.5.1 Magnetic Drum Storage Pulse Doppler System 113
6.5.2 Film Storage Electronic System 117
6.5.3 Electron Storage Tube System 119
6.6 Conclusions and Recommendations 123

CONTENTS (continued)

7. Measurement Errors 126

7.1 Statistical Errors in Cross-Correlation and Cross-Spectral Estimation 126

7.1.1 Cross-Correlation Estimation 127

7.1.2 Cross-Spectral Estimation 132

7.1.3 R-C Averaging 135

7.2 Phase Errors 137

7.2.1 Ordinary Spectral Density and Auto-correlation Measurements 137

7.2.2 Cross-Correlation and Cross-Spectral Density Measurements 147

7.3 Corrections for Finite Size Transducers in the Measurement of Boundary Layer Pressure Fluctuations . . 172

7.4 Magnetic Tape Recorder Dynamic Phase Errors 183

7.4.1 Tape Motion Equations 183

7.4.2 Analytical Examples 188

References 201

LIST OF FIGURES

Figure 2-1.	Plot of Two Sets of Identical Data with Different Arrangements	4
Figure 3-1.	Block Diagram of an Analog Correlation Analyzer	8
Figure 3-2.	Block Diagram of a Real Time Correlation Analyzer	9
Figure 3-3.	Magnetic Drum Variable Time Delay Mechanisms	10
Figure 3-4.	Magnetic Loop Delay Technique	12
Figure 3-5.	Logarithmic Multiplier	14
Figure 3-6.	Hall Effect Multiplier	15
Figure 3-7.	Quarter Square Multiplier	16
Figure 3-8.	Time Division Multiplier	16
Figure 3-9.	Waveforms in a Time Division Multiplier	18
Figure 3-10.	Operational Amplifier Integrator	20
Figure 3-11.	Block Diagram of an Analog Cross-Spectral Density Analyzer	24
Figure 3-12.	Real Time Cross-Spectral Density Analyzer	25
Figure 3-13.	90^{o} Phase Shift Circuit for Narrowband Signals	26
Figure 3-14.	Cross-Spectral Density Analyzer Using Heterodyne Techniques	28
Figure 3-15.	Zero IF Frequency Cross Spectral Density Analyzer.	30
Figure 3-16.	Cross Spectral Density Analyzer Using Direct Fourier Transform Techniques	33
Figure 4-1.	Binary Storage Locations for Fourier Transform	54
Figure 4-2.	Spectral Analysis Filter	60
Figure 4-3.	Illustration of Decimation	66
Figure 4-4.	Filter Frequency Response of Sin x/x Form	68
Figure 4-5.	Frequency Response Function of Single Tuned Filter	72
Figure 4-6.	Sequence of Single Tuned Filters for Spectral Analysis	72

LIST OF FIGURES (continued)

Figure 4-7.	Illustration of Variable Bandwidth Procedure	75
Figure 4-8.	Filter Frequency Responses for 1/3 Octave Spectral Analysis	77
Figure 4-9.	Computational Time Estimates as Functions of Number of Data Points m = 1000, B_a = 1/1000 Δt	81
Figure 4-10.	Computational Time Estimates as Functions of Number of Data Points, m = 500, B_a = 1/500 Δt	82
Figure 4-11.	Computational Time Estimates as Functions of Number of Data Points m = 100, B_a = 1/100 Δt	83
Figure 5-1.	Delta Modulation Pulse Code	91
Figure 5-2.	Delta Modulation Correlator	93
Figure 5-3.	Actual Delta Modulation Correlator Block Diagram	93
Figure 5-4.	Analog Sampling Hybrid Correlator	94
Figure 6-1.	Functional Diagram of a Magneto-Optical Playback System	106
Figure 6-2.	Functional Diagram of a Cathode Ray Tube	109
Figure 6-3.	Block Diagram of a Pulse Doppler System	114
Figure 6-4.	Block Diagram of a Film Storage System	117
Figure 6-5.	Block Diagram of a Film Readout System	118
Figure 6-6.	Block Diagram of a Correlation System	118
Figure 6-7.	Functional Diagram of a Correlator Using an Electron Storage Tube	121
Figure 7-1.	Transducers	172
Figure 7-2.	Attenuation Factor $F(\omega)$ for Circular Transducer of Uniform Sensitivity	177
Figure 7-3.	Attenuation Factor $F(\omega)$ for Circular Transducer of Uniform Sensitivity	178
Figure 7-4.	Transducer Sensitivity versus Radius	180
Figure 7-5.	Correction Factor for Finite Size Transducer	181

LIST OF FIGURES (continued)

Figure 7-6.	Sensitivity of B & K 4134 Microphone	182
Figure 7-7.	Data Related Phase Error Term for Sinusoidal Data . . .	190
Figure 7-8.	Spectral Density of Ideally Low Passed White Gaussian Noise.	191
Figure 7-9.	Data Related Phase Error for Data with an Ideal Low Passed Spectral Density	192
Figure 7-10.	Gaussian Shaped Spectral Density	193
Figure 7-11.	Data Related Phase Error Term for Data with Gaussian Shaped Spectral Density	194
Figure 7-12.	Flutter Related Phase Error for Sinusoidal Flutter	196
Figure 7-13.	Ideal Bandpassed White Gaussian Flutter Spectrum	197
Figure 7-14.	Dynamic Phase Error Spectrum Corresponding to Ideal Bandpassed Flutter Spectrum	198
Figure 7-15.	Flutter Related Phase Error Term for White Band-limited Flutter	200

1. INTRODUCTION

The purpose of this study is to investigate those factors contributing to inaccuracies in the measurement of cross-correlation and cross-spectral density functions of stationary dynamic pressure data. All aspects of the measurement problem from the transducer through the analyzing equipment are considered. In addition to the basic data acquisition equipment, analog, digital and hybrid techniques for the reduction of the data are investigated. Further, a survey of those techniques used in the radar and communication fields is made to determine which techniques are potentially applicable to the measurement of dynamic pressure data. Formulas are developed to relate the errors in actual hardware to the inaccuracy in the final measured cross-spectral density and cross-correlation functions.

This study was conducted to determine the ability of commercially available hardware to satisfy two specific requirements. In both cases dynamic pressure data are being measured. The first set of data is obtained from wind tunnel tests. For this type of test,50 channels of data each having a frequency response from 30 Hz to 100 KHz is required. The second set of data is obtained from flight tests. For this type of test,250 channels of 3 Hz to 10 KHz data are required. Although this study is conducted to specifically satisfy the above two data channel requirements, most of the results are presented in sufficiently general terms to be useful for vastly different applications.

Section 2 contains a brief mathematical description of the cross-correlation and cross-spectral density functions which are the object of this analysis. Section 3 describes analog techniques that are used to compute both the cross-spectral density and the cross-correlation function. Section 4 describes digital techniques that are used to compute cross-correlation and cross-spectral density functions. Section 5 contains a description of

hybrid computer techniques that are used to compute cross-correlation and cross-spectral density functions. Section 6 reports a survey of the radar and communication fields and a description of the techniques used in those fields which presently appear to be feasible for application to the measurement of dynamic pressure data. Section 7 describes the measurement errors. This section discusses those errors distributed throughout the entire measurement system - transducer errors, signal conditioning errors, transmission errors, recording errors, and data reduction errors. Both hardware related errors and statistical uncertainty errors are covered.

2. DEFINITIONS OF CORRELATION AND SPECTRAL DENSITY FUNCTIONS

Correlation, statistically speaking, is a measure of linear relationship between two variables. Random variables are of primary interest when this concept is employed, although the idea is also of use for deterministic functions such as sinusoids.

In basic statistics, the correlation between two variables x and y is defined as the average cross product between these two variables. That this is a measure of average linear relationship is easily seen from the following simple sets of data.

Case 1			Case 2		
x	y	xy	x	y	xy
-2	-3	6	-2	0	0
-2	-1	2	-2	0	0
-1	0	0	-1	-1	1
0	0	0	0	-3	0
+1	+1	1	+1	+3	3
+2	+3	6	+2	+1	2
	$\Sigma xy =$	15		$\Sigma xy =$	6

The two individual sets of data are identical, but in Case 1 the arrangement roughly matches small numbers with small and large numbers with large. In the second case, the set of y data is randomly scrambled to remove this relationship. One can clearly see the effect of the linear relation on the size of the sum of the cross products. The amount of linear relationship in the two above sets are further illustrated in Figure 2. 1. Notice that the x-y relation in Case 1 is much closer to being a perfectly linear relation than that in Case 2.

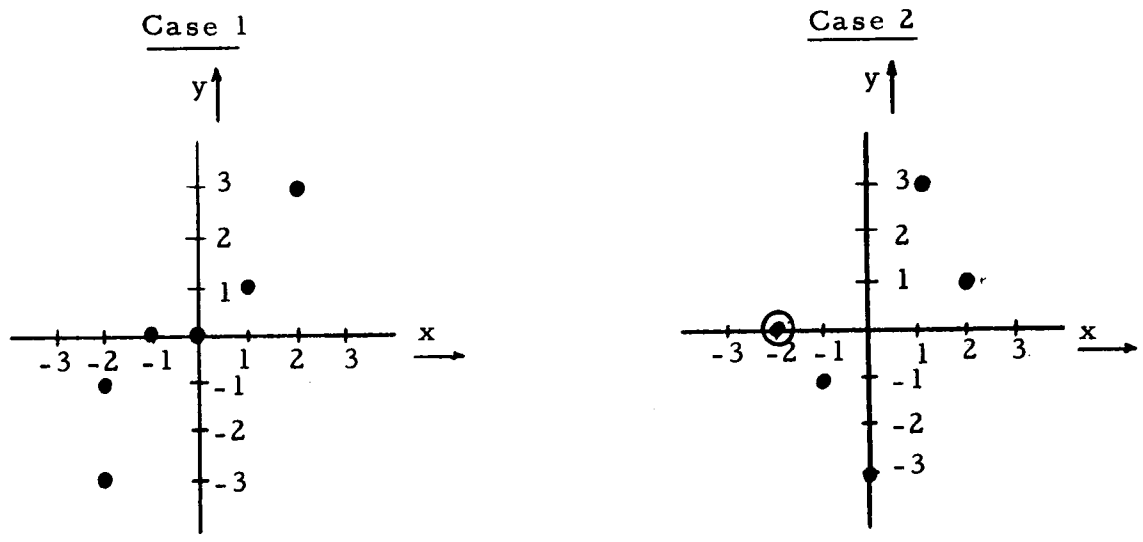


Figure 2-1. Plot of Two Sets of Identical Data with Different Arrangements

The type of correlation of interest for this report arises when x and y are variables taken as a function of time. When this is the case, there are two possible types of correlation functions, either a cross-correlation function in the case of two variables x(t) and y(t), or an autocorrelation function in the case of just one variable. That is, the time average of cross products of the variable x with the variable y where the results are presented as a function of time delay (τ). The variable y is delayed with respect to the variable x. In statistical terms, for stationary processes, the definition is

$$E[x(t)\, y(t+\tau)] = R_{xy}(\tau) \qquad (2.1)$$

where E[ ] denotes the expected (average) value of the quantity in brackets. For convenience, it is assumed here that the mean values of these processes are either originally zero or that they have been subtracted out. This causes no loss in generality. Many transducers have no DC response and hence a zero mean value time series is automatically obtained. In general, a computer

program to compute correlation functions and spectral density functions will do certain preprocessing of the data including subtracting out of the mean value. Thus, by the time the data is presented to the correlation and spectrum computation program, it will be a zero mean process.

For zero mean stationary processes, $R_{xy}(\tau)$ is a covariance function which in this case is identical to the correlation function. The term covariance arises from the fact that this quantity is a measure of the covariability of the two variables x(t) and y(t). The term covariance is rigidly restricted to the average cross product of two variables which have zero mean values. Thus, if the mean is non-zero, it must be computed and subtracted out. In classical statistics, the discussion of correlation is usually restricted to correlation coefficients which are defined as covariances normalized to lie in the range minus one to plus one.

In the case where x and y are continuous functions of time, Eq. (2.1) is defined by an integration. The mathematical expression for the sample estimate from finite length sample records is, for the continuous case

$$\hat{R}_{xy}(\tau) = \frac{1}{T-\tau}\int_0^{T-\tau} x(t)\, y(t+\tau)\, dt, \qquad |\tau| \leq \tau_{max} < T \tag{2.2}$$

where the hat "^" denotes the estimate. The quantity τ_{max} defines the range of time delay values from which the correlation function is computed. For the discrete case, the expression is

$$\hat{R}_{xy}(r\Delta\tau) = \frac{1}{N-r}\sum_{i=1}^{N-r} x(i\Delta t)\, y(i\Delta t + r\Delta\tau), \quad |r| = 0, 1, \ldots, m \tag{2.3}$$

where

$$N\Delta t = T, \quad m\Delta t = \tau_{max}$$

$$\Delta\tau = \Delta t \text{ (usually)}$$

As one can see from an inspection of Eq. (2.3), the primary computational procedure is that of multiplying x times y and then summing over the N observed data points of the finite length record. This procedure is then repeated for all values of r. In practice the maximum lag number m should almost never exceed half the value of the total number of observations N, and usually is restricted to about one-tenth or less of the size of N. For the autocorrelation function, one sets $y(t) = x(t)$, and performs the identical computations.

Since the frequency composition of a random variable is often of interest, generalized harmonic analysis of a random variable is often performed. When one is interested in this type of an analysis, the power spectral density function of a single record or the cross-spectral density function of a pair of records is the parameter of interest. The power spectral density function is a characterization of the random variable x(t) in terms of its variance (mean square value) or "power" as a function of frequency. That is, the process is decomposed into its various frequency components in order to determine the amount of power or variance in each frequency band. The name is then self-explanatory in that a characterization of a given variable as a function of frequency is termed a spectrum. The variance of a signal can be shown to be essentially equivalent to power (up to a scale factor). By definition, the power spectral density function is a density function since it is normalized with respect to frequency. Hence, the term power spectral density function. The cross-spectral density function is similar and may be thought of as the covariance between a pair of random variables x and y as a function of frequency.

In practice, one cannot obtain values of these functions at discrete frequency points; instead, one obtains average values within small frequency bands of width B_a. The usual analog method of obtaining a power spectral density function is to filter a record within a small bandwidth B_a, then square and average the filter output. This gives the mean square value within the

small frequency band B_a centered about a frequency f. This process is repeated for all frequency values f which are of interest.

With the analogies of the spectra to variance and covariance, it is possibly not too surprising that the power and cross-spectral density function can be obtained from auto and cross-correlation functions. By a famous formula termed the Wiener-Khinchine relation, the power spectral density function and the cross-spectral density function may be transformed into the autocorrelation function and the cross-correlation function respectively. This formula gives the spectral density function as the Fourier transform of the corresponding correlation function. In equation form, this is

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)\, e^{-j 2\pi f \tau}\, d\tau \quad , \quad -\infty < f < \infty \qquad (2.4)$$

$$= \int_{-\infty}^{\infty} R_{xy}(\tau) \left[\cos 2\pi f\tau - j \sin 2\pi f\tau\right] d\tau$$

where $j = \sqrt{-1}$. This may be intuitively pictured as picking out the second order periodicities by "beating" the correlation function with in phase and 90^{o} out of phase sinusoids.

The cross-spectral density function is a complex number and hence consists of two components. The real part is termed the cospectral density function and the imaginary part termed the quadrature spectral density function denoted by $C_{xy}(f)$ and $Q_{xy}(f)$, respectively. For the ordinary power spectral density function, $R_{xy}(\tau)$ is replaced by $R_{xx}(\tau)$. Then $S_{xx}(f)$ is a real-valued function since $R_{xx}(\tau)$ is an even function and hence the sine portion of the Fourier transform is zero. The importance of Eq. (2.4) lies in the fact that if the power spectral density function is of interest in a given application, it may also be obtained via the correlation function with the additional computational step of a Fourier transform. A broad discussion of these matters and other basic requirements for measuring and analyzing random data appear in Reference 1.

3. ANALOG DATA REDUCTION TECHNIQUES

This section contains a summary of the techniques used in analog instruments for the computation of correlation functions and cross-spectral density functions. In addition to describing the basic computational techniques, various practical methods of implementing the required functional operations are discussed.

3.1 CORRELATORS

Since the equation for computing a correlation function is as follows,

$$R_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, y(t + \tau)\, dt \qquad (3.1)$$

one of the two signals under analysis must be delayed with respect to the other and then the two signals must be multiplied and their product integrated. Since one really only estimates the correlation function, the limiting process is dropped. The uncertainty errors associated with finite record lengths and bandwidths are described in Section 7. A block diagram of a typical analog correlator is shown in Figure 3-1.

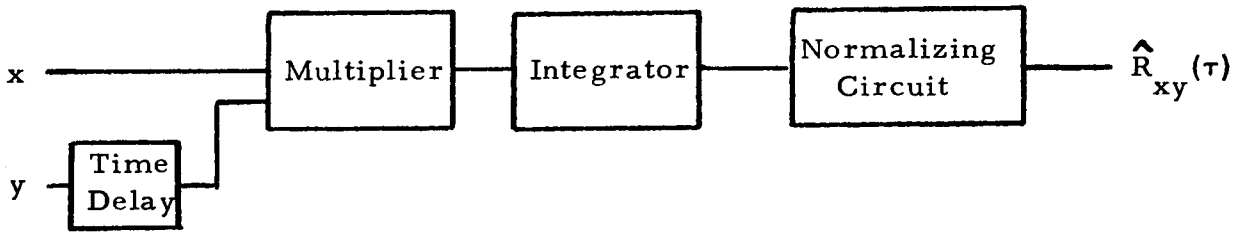


Figure 3-1. Block Diagram of an Analog Correlation Analyzer

Two basic versions of this type of correlator are possible. The primary difference is in the manner of obtaining $R_{xy}(\tau)$ for various values of time delay (τ). In the first, and most common type, the value of time delay is scanned across the time delay range of interest. Essentially, a complete analysis is performed for one value of τ and then a new value of τ is set into the delay mechanism and the data is repeated and so on until the complete time delay range of interest is analyzed. The time delay can be stepped in discrete increments or can be continuously swept.

The second basic type of analog correlation analyzer consists of a number of parallel circuits. A separate circuit is required for each value of τ. See Figure 3-2.

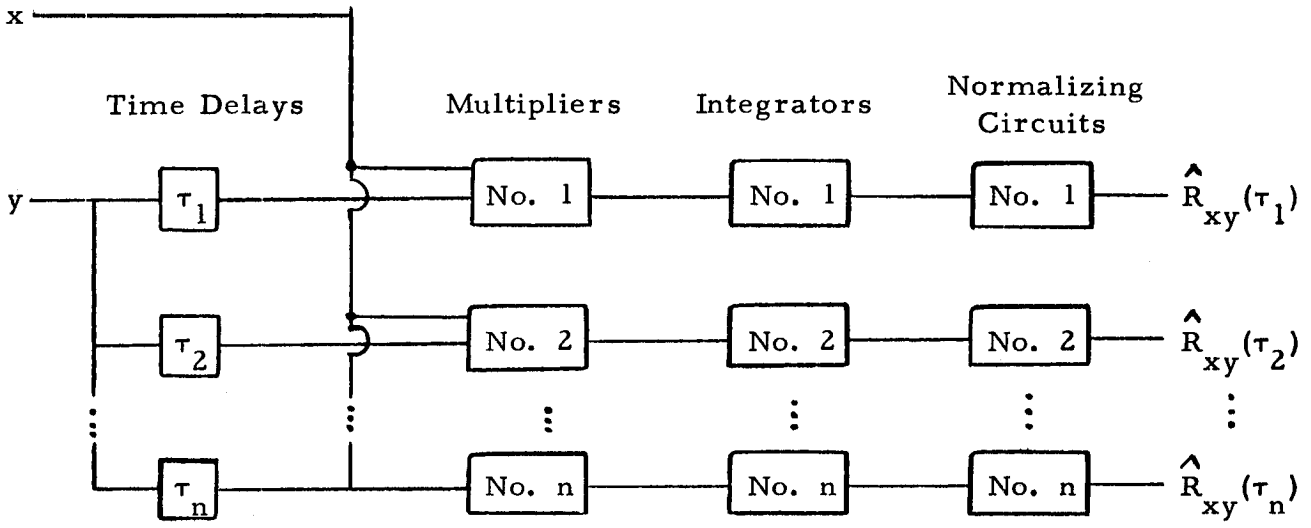


Figure 3-2 Block Diagram of a Real Time Correlation Analyzer

These parallel analyzers are frequently termed real time analyzers since they have the capability to perform a complete analysis as the data is being generated. No storage of the input data is required.

Unfortunately, two of the three operations that must be performed on the data to compute correlation functions are difficult to perform with electrical analog equipment. These operations are the time delay and the multiplication operations.

3.1.1 Time Delay Mechanisms

The most successful of the analog time delays are based on mechanical devices. For example, if both data channels are recorded on separate tracks of a magnetic drum, the time delay can be obtained by moving the position of one reproduce head relative to the other (Reference 2.). See Figure 3-3. Or, in a similar scheme, only one data channel is recorded on the magnetic drum. This is the channel to be delayed. The position of the reproduce head is varied to obtain the desired time delay. Normally, separate record and reproduce heads are used so that there is some minimum separation required between the two magnetic heads because of the mechanical interference. Since this separation represents a limitation on the minimum time delay value, it is necessary to delay the other data channel by a fixed amount equal to or greater than the minimum delay value of the drum.

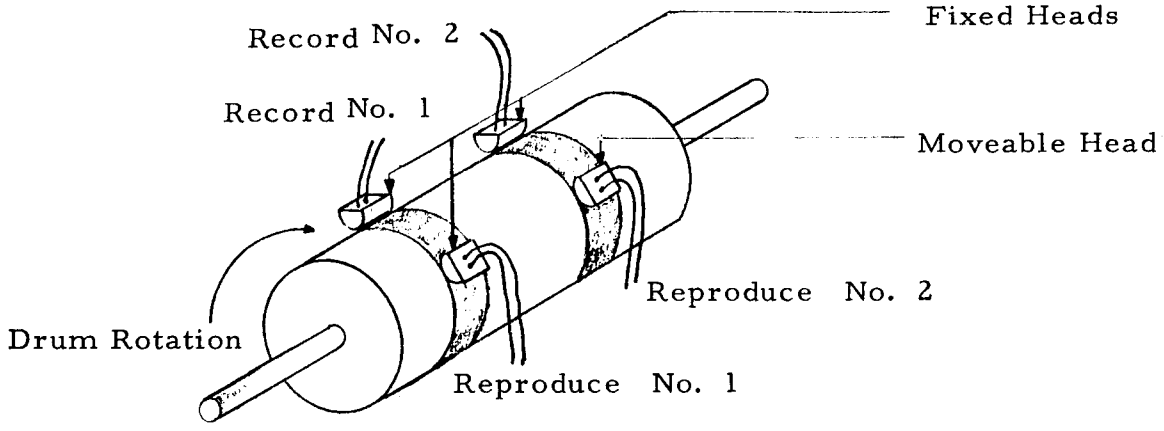


Figure 3-3. Magnetic Drum Variable Time Delay Mechanisms

Another version of the magnetic drum delay mechanism consists of one recording track with fixed record and reproduce heads. The variation in time delay is obtained by varying the rotational speed of the drum (Reference 3). As in the preceding case, a fixed delay is required in the other data channel to compensate for the delay associated with the physical separation of the record and reproduce heads.

All of the above schemes for creating a time delay can also be implemented by using a magnetic tape recorder in place of the drum. The main advantage of using tape is that magnetic tape is in contact with the magnetic heads; hence, the mechanical mechanisms required, although complex, are not as complex as those required with a drum where the head must be maintained very close to but not in contact with the magnetic surface. The primary disadvantage of using magnetic tape is that its velocity cannot be controlled as accurately as that of a drum. For example, temperature, humidity, and tension variations all cause the tape to stretch in ranges where these effects would be negligible for a drum.

One magnetic tape recorder that has been specifically designed to provide time delays has two separate capstans and a storage bin (Reference 4). The signals to be correlated are recorded on separate tracks of the tape. To provide a fixed value of time delay, one capstan is engaged and tape is dumped into the storage bin until the desired value of time delay is obtained. Then the second capstan is engaged and run at the same speed as the first capstan. In this manner a fixed time delay is created. If it is desired to continuously vary the time delay, the second capstan is run at a slightly different speed than the first. The excess tape generated by this speed differential is temporarily stored in the storage bin. Two separate record/reproduce head stacks are required to obtain zero time delay.

Another mechanism that has been successfully used to generate time delays operates on the principle of overlaying permanent records of the time

histories of the two signals to be correlated. To obtain a time delay one record is physically displaced along the time axis relative to the other. This technique is used in optical correlators (Reference 5). Here the data is recorded in the form of intensity modulation on photographic film, each signal to be correlated on a separate film. The same scan rate (inches/sec) of the light trace is used in all traces so that when the films are overlayed such that the starting points of both records coincide, the correlation function value for $\tau = 0$ can be generated. By moving one film relative to the other x inches a time delay of x divided by the scan rate is generated.

This same technique is also used in a magnetic correlator (Reference 6) The two signals to be correlated are recorded on separate loops of magnetic tape. One loop is positioned with the oxide side of the tape against one side of a long flux sensitive reproduce head. The other tape loop is positioned with its oxide side against the other side of the flux sensitive head. See Figure 3-4. Displacement of the second loop along the magnetic head relative to the first loop creates the time delay.

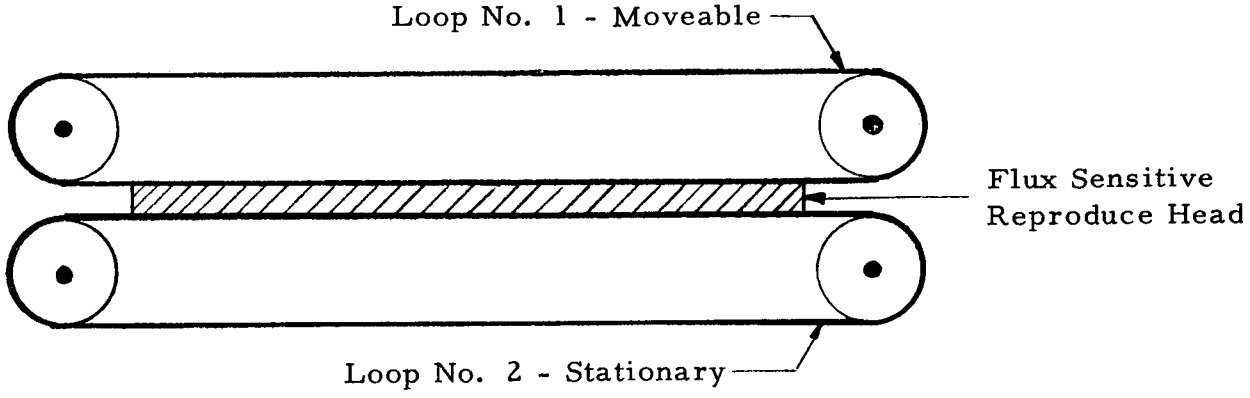


Figure 3-4. Magnetic Loop Delay Technique

Electronic delay lines have not found much usage as a source of variable delay for acoustic data correlators. These devices usually act as low-pass or bandpass filters so that direct application of a delay line to a signal extending from 3 Hz to 100 KHz is not too practical for moderate values of delay (0.1 sec). Modulation of a higher frequency carrier is necessary to obtain a 100 KHz bandwidth. Also it is usually necessary to tap the delay line to obtain the different values of delay. Providing even a moderate number of taps ($\sim$200) and correcting for the difference in insertion loss at each of the taps, though feasible, is a bit unwieldy and the end result is one that does not have the operational flexibility of some of the other delay techniques.

3.1.2 Multiplication Mechanisms

A large number of electronic analog multiplication techniques have been devised over the years. For a discussion of most of these, see Reference 7. The techniques can be separated into general categories of special electronic tubes, special modulation formats, quarter square multipliers, Hall effect multipliers, time-division multipliers and logarithmic multipliers. The latter four techniques have most frequently been applied to acoustic type data analyzers.

Logarithmic multipliers simply operate on the principle

$$\log xy = \log x + \log y \qquad (3.2)$$

Each data channel is applied to a separate log converter, summed and the antilog is taken as shown in Figure 3-5.

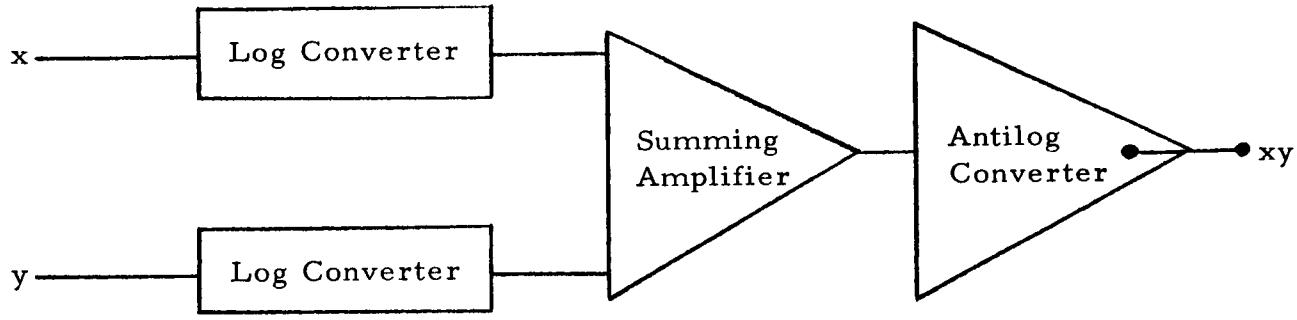


Figure 3-5. Logarithmic Multiplier

In general, this type of multiplier is not too satisfactory for correlation analysis since it operates in one quadrant only (x and y must both be positive).

The Hall effect multipliers are constructed from semiconductor material. The output of this device is proportioned to the product of the current applied to the semiconductor and the flux density applied in a perpendicular plane.

$$V_z = \frac{R_h I_x B_y}{t} \tag{3.3}$$

where

V_z = the output voltage

R_h = the Hall constant

I_x = a current proportional to x

B_y = a flux density proportional to y

t = the thickness of the semiconductor material

Figure 3-6 illustrates the basic operation of a Hall effect multiplier.

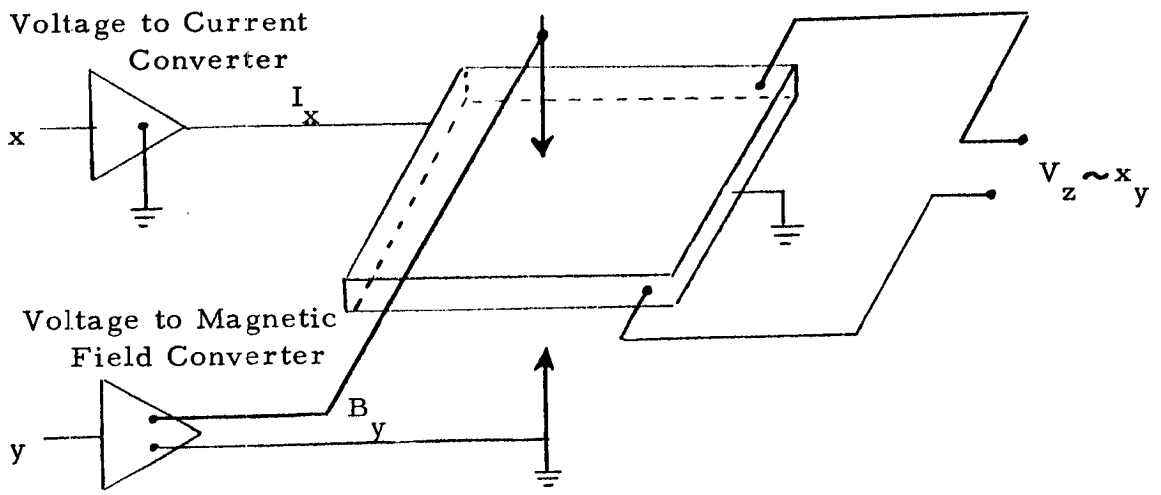


Figure 3-6. Hall Effect Multiplier

These devices presently have high frequency limitations on the order of 10 KHz.

The quarter-square and time division multipliers are the most commonly used for this specific application. Quarter square multiplication is based on the algebraic identity

$$xy = \frac{1}{4}\left[(x + y)^2 - (x - y)^2\right] \tag{3.4}$$

A block diagram of this type of multiplier is shown in Figure 3-7.

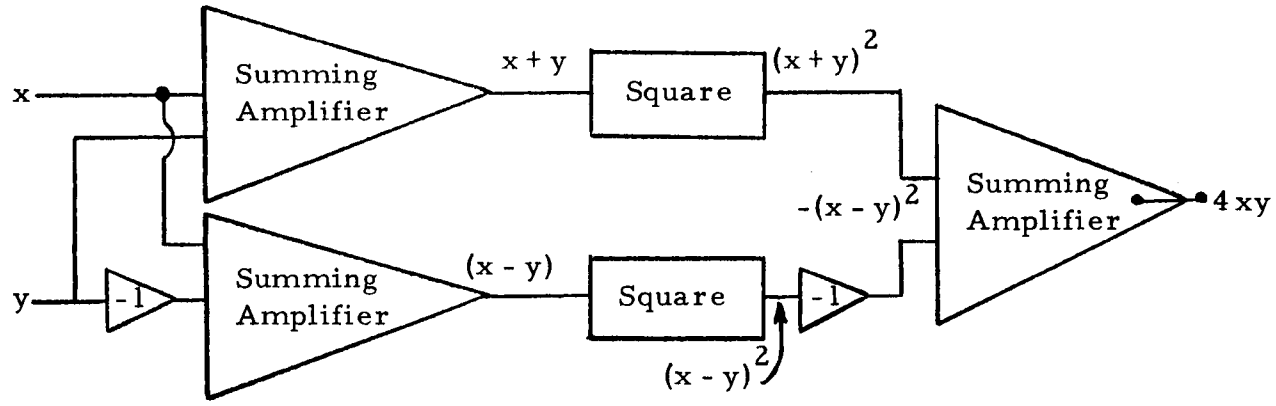


Figure 3-7. Quarter Square Multiplier

This technique is popular because a large amount of work has gone into the development of accurate squaring devices and because the other electronics required are simply amplifiers.

Time division multiplication is essentially a combination of pulse amplitude and pulse duration modulation. One of the two signals to be multiplied is passed through a pulse duration modulator to generate a series of constant amplitude pulses whose ratio of "on" to "off" time is a function of the instantaneous value of the data signal. The repetition rate of the pulses must be much higher than the highest data frequency. These pulses are then used to sample the second data signal. Analog samples are taken in the second channel. (This can be thought of as pulse amplitude modulation where the sample rate is not constant.) The sampled signal is then integrated to obtain the product since it is a function of the area in the sample pulses. The pulse width is proportional to the signal level in the first channel and the pulse height is proportional to the signal level in the second channel. The following figure is a block diagram of a time division multiplier.

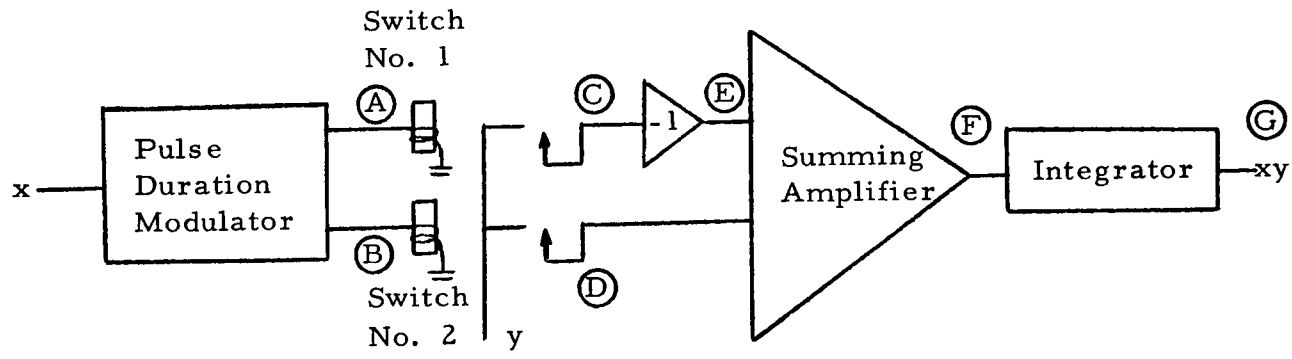


Figure 3-8. Time Division Multiplier

To further illustrate the operation of this multiplier consider Figure 3-9 where the waveforms throughout Figure 3-8 are shown. As an example, assume that x is a positive dc voltage, and y is a positive dc voltage. The positive voltage into the PDM circuit causes the switch "on time" (t_1) at point (A) to be less than the switch "off time" (t_2). Conversely, this positive value of x causes the switch "on time" at point (B) to be greater than the switch "off time." At point (C) the value of y is seen during the time switch No. 1 is operated and is zero the rest of the time. At point (D) the value of y is seen when switch No. 2 is on and zero is seen when it is off. Note that when y is not seen at point (C), it is seen at point (D) and vice versa. Point (E) contains an inverted version of point (C). The pulses now swing in voltage from zero to minus y instead of zero to plus y. Also note that the time history at point (E) is identical to that at point (D) except that it has been effectively biased by an amount of minus y. This step is taken to get a zero output when x = 0, ($t_1 = t_2$ in this situation). The voltage at point (F) is the sum of those occurring at (D) and (E). This swings between plus and minus y. The voltage at point (G) is the mean value of that at point (F), a dc value proportional to the product xy in this example.

3.1.3 Integration Mechanisms

Integration is easily performed with electronic circuits. Expressed in Laplace transforms (Reference 8)

$$\int_0^t f(t)\, dt = \left(\frac{1}{s}\right) F(s) \qquad (3.5)$$

where F(s) = the Laplace transform of f(t).

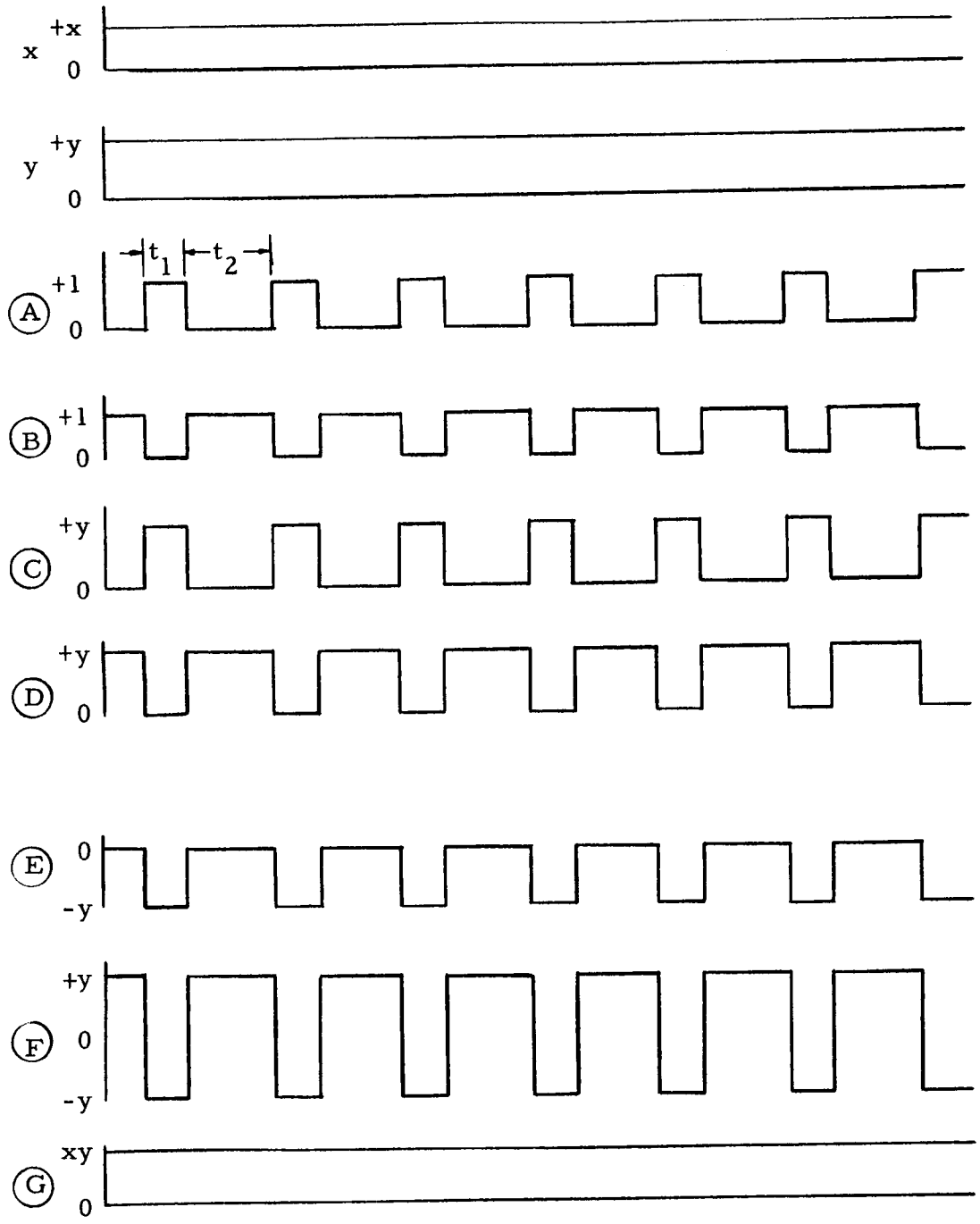


Figure 3-9. Waveforms in a Time Division Multiplier

In practice the (1/s) transfer function is approximated by an $\frac{\alpha}{s + \alpha}$ transfer function where α is the reciprocal of the RC time constant of a simple RC lowpass filter circuit. The error in the approximation can be determined from the difference of the inverse Laplace transform of the perfect integrator and the RC approximate integrator.

$$\mathcal{L}^{-1}\left[\frac{1}{s}\right] = 1, \text{ or a unit step function at } t = 0 \tag{3.6}$$

$$\mathcal{L}^{-1}\left[\frac{\alpha}{s + \alpha}\right] = \alpha\, e^{-\alpha t} \tag{3.7}$$

Ignoring the scale factor

$$e = \left(\frac{1 - e^{-t/RC}}{1}\right) 100 \quad , \quad t \geq 0 \tag{3.8}$$

where e = the normalized percentage error. For example, if a 5% error is permissible, one finds that the maximum integration time, t_{max} equals 0.051 RC.

To obtain long time constants without the serious signal attenuation attendant with simple RC integrators, operational amplifier integrators are used. The circuit of this type of amplifier is shown in Figure 3-10.

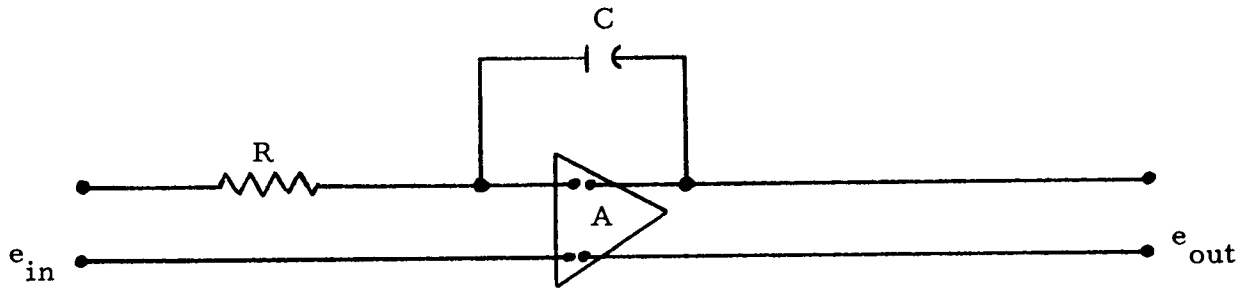


Figure 3-10. Operational Amplifier Integrator

Ignoring parasitics in the computing elements and imperfections in the operational amplifier, other than finite gain (A), it can be readily shown that the transfer function of the above circuit is

$$\frac{e_{out}}{e_{in}} = \left(\frac{-A}{A+1}\right)\left(\frac{1}{RC}\right)\left(\frac{1}{s + \frac{1}{RC(A-1)}}\right) \tag{3.9}$$

If the magnitude of the gain is quite large compared to unity, Eq. (3.7) can be rewritten

$$\frac{e_{out}}{e_{in}} \approx \left(\frac{-1}{RC}\right)\left(\frac{1}{s + \frac{1}{ARC}}\right) \tag{3.10}$$

In the time domain

$$\frac{e_{out}}{e_{in}} \approx \left(-\frac{1}{RC}\right)\left(e^{-t/ARC}\right) \tag{3.11}$$

Comparing Eqs. (3.7) and (3.11), it can be seen that for a given RC product, the operational amplifier increases the effective RC time constant by approximately the amount of the gain of the amplifier while maintaining the signal attenuation constant, although inverting it in phase.

3.2 CROSS-SPECTRAL DENSITY ANALYZERS

The classical definition of cross-spectral density (see Reference 9) is

$$S_{xy}(f) = \lim_{T\to\infty} \frac{1}{2T}\left[\int_{-T}^{T} x(t)\, e^{-j2\pi ft}\, dt\right]\left[\int_{-T}^{T} y(t)\, e^{j2\pi ft}\, dt\right] \tag{3.12}$$

where

$S_{xy}(f)$ = the two-sided cross-spectral density function

$x(t)$ = one of the data signals to be analyzed

$y(t)$ = the other data signal to be analyzed

For physically realizable applications, one is interested in the one-sided, measurable cross-spectral density function, $G_{xy}(f)$.

$$G_{xy}(f) = 2\,S_{xy}(f) \quad ; \quad f \geq 0$$
$$= 0 \quad ; \quad f < 0 \tag{3.13}$$

The cross-spectral density function is frequently described in terms of its real and imaginary components.

$$G_{xy}(f) = C_{xy}(f) - j\,Q_{xy}(f) \tag{3.14}$$

$$C_{xy}(f) = \lim_{T\to\infty,\, B\to 0} \frac{1}{2BT} \int_{-T}^{T} x(f, B, t)\, y(f, B, t)\, dt \tag{3.15}$$

$$Q_{xy}(f) = \lim_{T\to\infty,\, B\to 0} \frac{1}{2BT} \int_{-T}^{T} x(f, B, t)\, y^{o}(f, B, t)\, dt \tag{3.16}$$

where

$C_{xy}(f)$ = the cospectral density

$Q_{xy}(f)$ = the quadrature spectral density

B = the bandwidth of the analyzer filter

f = the center frequency of the analyzer filter

$x(f, B, t)$ = one of the data signals to be analyzed after it has been passed through a bandpass filter of width B, and center frequency f

$y(f, B, t)$ = the other data signal passed through an identical filter

$y^{o}(f, B, t) = y(f, B, t)$ shifted in phase by 90^{o}

The cross-spectral density can also be described in polar form.

$$G_{xy}(f) = \left| G_{xy}(f) \right| e^{-j \phi_{xy}(f)} \tag{3.17}$$

where

$$\left| G_{xy}(f) \right| = \text{the magnitude factor} = \sqrt{C_{xy}^{2}(f) + Q_{xy}^{2}(f)} \tag{3.18}$$

$$\phi_{xy}(f) = \text{the phase factor} = \tan^{-1}\left[\frac{Q_{xy}(f)}{C_{xy}(f)}\right] \tag{3.19}$$

The equations normally implemented in analog cross-spectral density analyzers are those shown in Eqs. (3.15) and (3.16). Since one can only estimate the values, the limiting processes are deleted.

A block diagram of a typical analog cross-spectral density analyzer is shown in Figure 3-11.

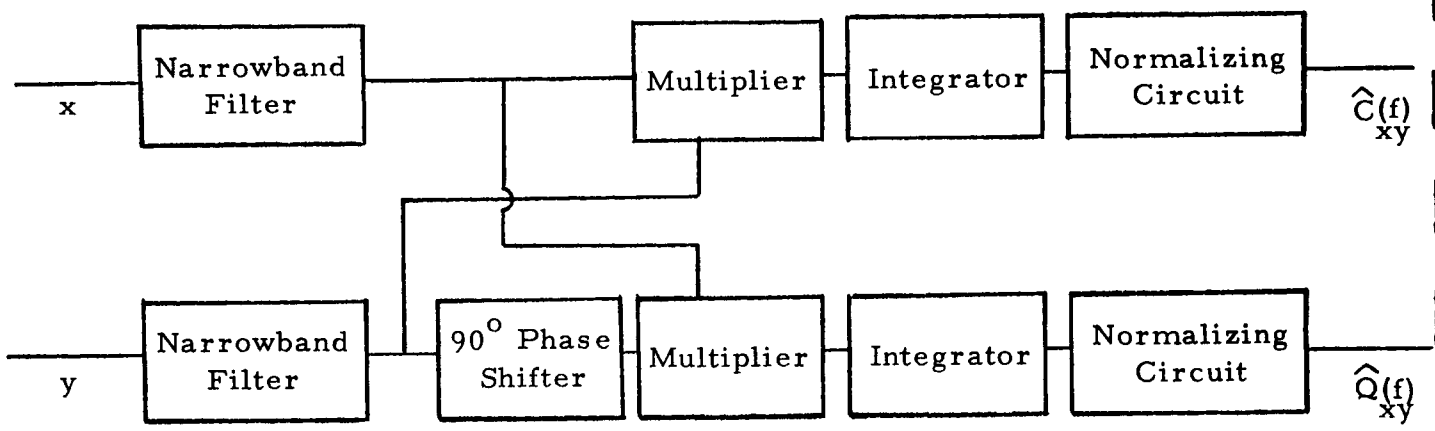


Figure 3-11. Block Diagram of an Analog Cross-Spectral Density Analyzer

The x(t) data signal is passed through a narrowband bandpass filter. The y(t) data signal is passed through an identical filter. To obtain the co-spectral density, the two filter outputs are multiplied and integrated. To obtain the quadrature spectrum, the filtered y(t) signal is shifted 90^{o} in phase and multiplied by the filtered x(t) signal. This product is then integrated.

As with the analog correlator, two basic versions are possible. A single channel analyzer as shown in Figure 3-11 whose center frequency is scanned over the frequency range of interest or a real time analyzer, as shown in Figure 3-12 with contiguous filters covering the frequency range of interest.

3.2.1 90^{o} Phase Shift Mechanisms

Referring to Figure 3-11, the analog computational steps required to compute cross-spectral density are narrowband bandpass filtering, 90^{o} phase shifting, multiplication, and integration. Analog techniques for performing multiplication and integration were previously discussed in Section 3.1 and will not be repeated here. The 90^{o} phase shift operation can

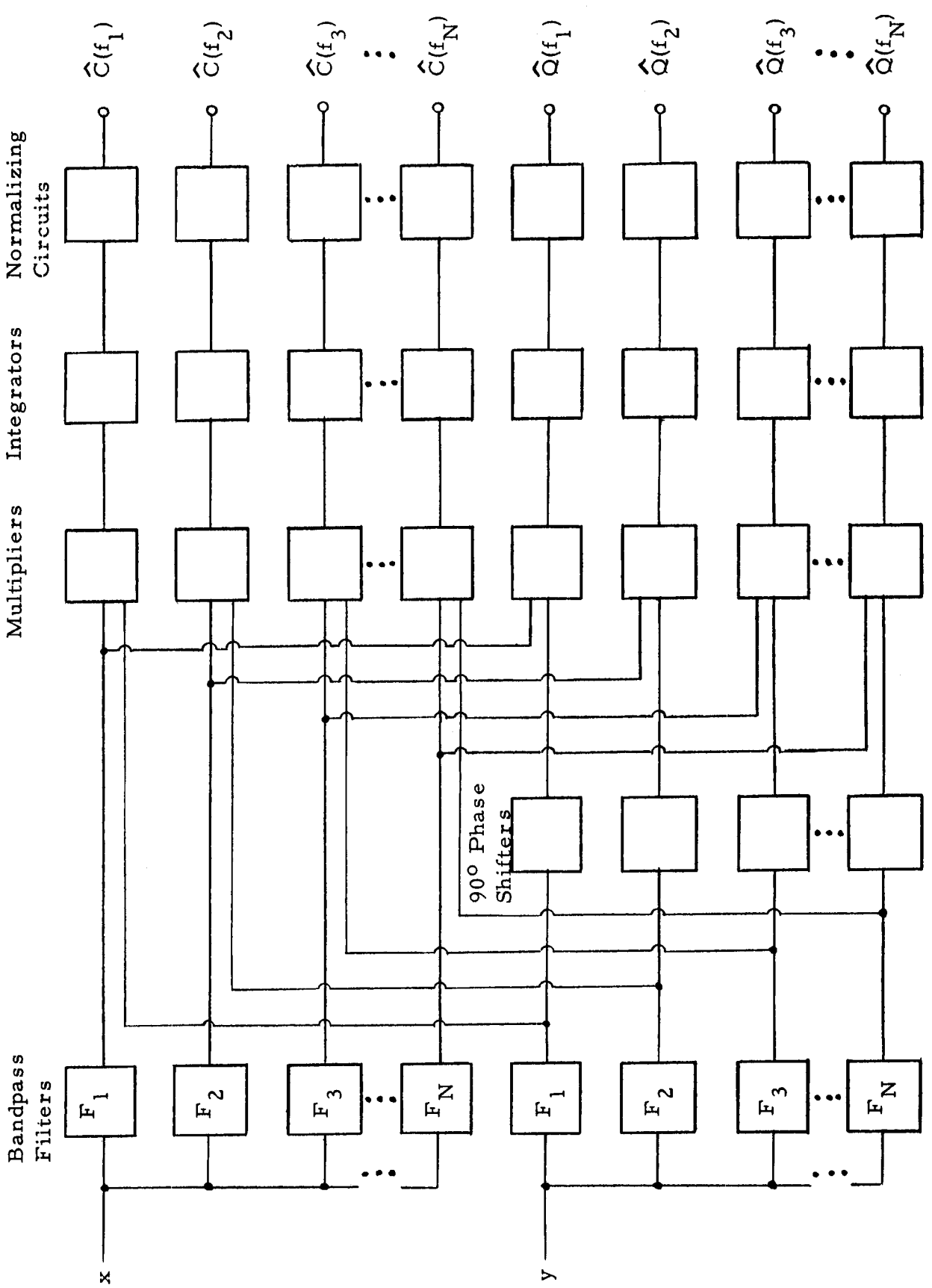


Figure 3-12. Real Time Cross-Spectral Density Analyzer

be implemented with a simple RC circuit as long as the bandwidth of the bandpass filter is much smaller than the center frequency of the filter. The RC circuit is shown in Figure 3-13.

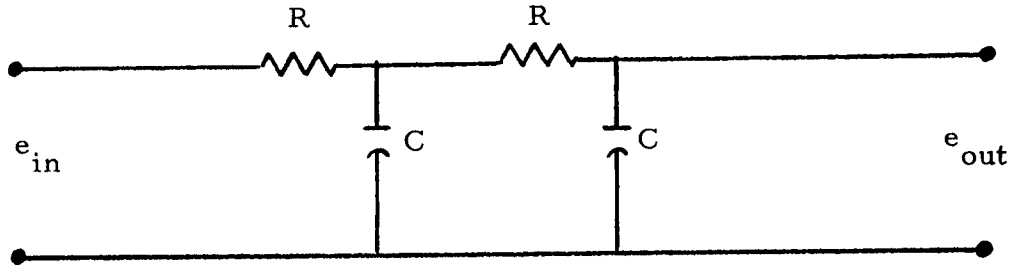


Figure 3-13. 90^{o} Phase Shift Circuit for Narrowband Signals

The input to output relation of this circuit is given in Eq. (3.20).

$$\frac{e_{out}}{e_{in}} = \left[\frac{1}{\left\{(\omega RC)^2 - 1\right\}^2 + (3\omega RC)^2}\right]\left[\left\{1 - (\omega RC)^2\right\} - j(3\omega RC)\right] \qquad (3.20)$$

The phase shift as a function of frequency is

$$\phi(\omega) = \tan^{-1}\left[\frac{-3\omega RC}{1 - (\omega RC)^2}\right] \qquad (3.21)$$

At $\omega RC = 1$, a 90^{o} phase shift between the input and output voltages is obtained. As long as $0.97 \leq \omega RC \leq 1.03$, there is about $\pm 1^{o}$ phase error, or less, over the data passband.

When the data is not directly filtered, the 90^{o} phase shift operation is most conveniently performed by simultaneously multiplying the data by sine and cosine waves. This operation is explained in detail in the following subsection on filtering mechanisms.

3.2.2 Filter Mechanisms

There are numerous manners in which bandpass filters can be constructed and a summary of all of these techniques is beyond the intent of this section. There are however, three basic filtering techniques used in analog cross-spectral density analyzers. The first technique is direct application of the filter(s) to the data signal, the second technique is heterodyning of the data signal to a different frequency range, and the third technique is based on direct Fourier transformation of the data.

Direct filtering of the data has the advantage of having the simplest circuitry (assuming the same filter is used in all three techniques). Thus, this technique is the most reliable and stable. However, it is also the least flexible.

The data can be heterodyned to a higher frequency for analysis with bandpass filters or to DC for analysis with lowpass filters. This latter technique is called zero frequency IF. Heterodyning of the data to a higher frequency range permits greater operational flexibility than direct filtering as well as the use of crystal and magnetostrictive filters. This principle is frequently used when one pair of "identical" filters is scanned through the frequency range of interest. By using one pair of "identical" filters tuned to a fixed frequency and slowly scanning the frequency of the local oscillator in the heterodyne circuits, the data frequency can be scanned past the fixed frequency filter which is effectively the same as scanning the filter through the data frequency range.

Figure 3-14 is a block diagram of a typical heterodyne filter circuit type of cross-spectral density analyzer. In most cases a balanced modulator is used to suppress the carrier frequency. Thus, only sum and difference frequencies (sidebands) are present on the output of the modulator. The bandpass filters are as closely matched in characteristics as possible. The filters work on only one of the sidebands. For example, assume that the filters have a center frequency of 100 KHz, and the data covers a bandwidth from 5 Hz to 5 KHz. By tuning the local oscillator to 99, 995 Hz, the 5 Hz data component is made to appear at the center frequency of the filter.

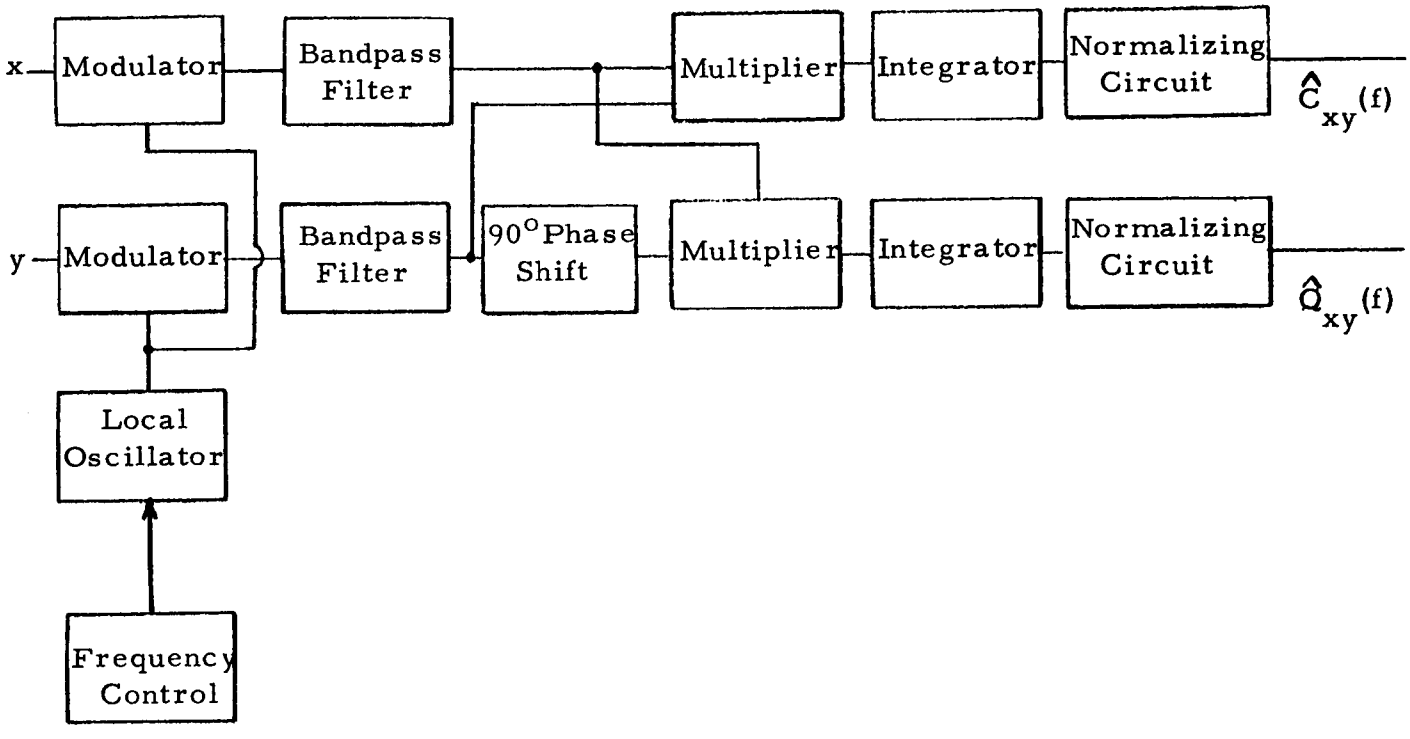


Figure 3-14. Cross-Spectral Density Analyzer Using Heterodyne Techniques

(Positive sideband component equals 100 KHz.) Similarly, by tuning the local oscillator to 95 KHz, the 5 KHz data components appear at the center frequency of the bandpass filter. Slowly varying the frequency of the local oscillator thus effectively causes the bandpass filter to be scanned through the data frequency range.

The zero frequency IF filtering technique offers the advantage of sharper filter rolloff characteristics and better phase matching for a given cost since the actual filter only needs to have a lowpass instead of a bandpass characteristic. This analyzer operates by simultaneously multiplying the inputs by sine and cosine waves as shown in Figure 3-15. The frequency of the oscillator is set to the frequency at which it is desired to filter the data. The output of the multiplier then contains sum and difference frequencies between the data and the oscillator frequencies. By lowpass filtering the data, only the difference frequencies are analyzed. The apparent bandwidth of the filter then extends from $f_0 - f_1$ to $f_0 + f_1$, where f_0 equals the oscillator frequency and f_1 equals the cutoff frequency of the lowpass filter. The required 90^o phase shift is performed by the simultaneous multiplication by sine and cosine terms.

One problem that has been encountered with this type of filtering is that a large amount of filter ripple occurs near its center frequency. This error occurs because the analysis record durations are finite while the actual center frequency is dc, which requires an infinite record length to avoid bias errors. (For a more detailed discussion of this error see Section 6.) The latest versions of analyzers employing zero frequency IF filtering have added additional circuitry that circumvents this problem.

Direct Fourier transformation of the data to obtain the cross-spectral density is based on implementation of Eq. (3.12). Rewriting this equation in terms of the estimate of the one sided cross-spectral density and replacing the exponential with trigonometric terms one obtains

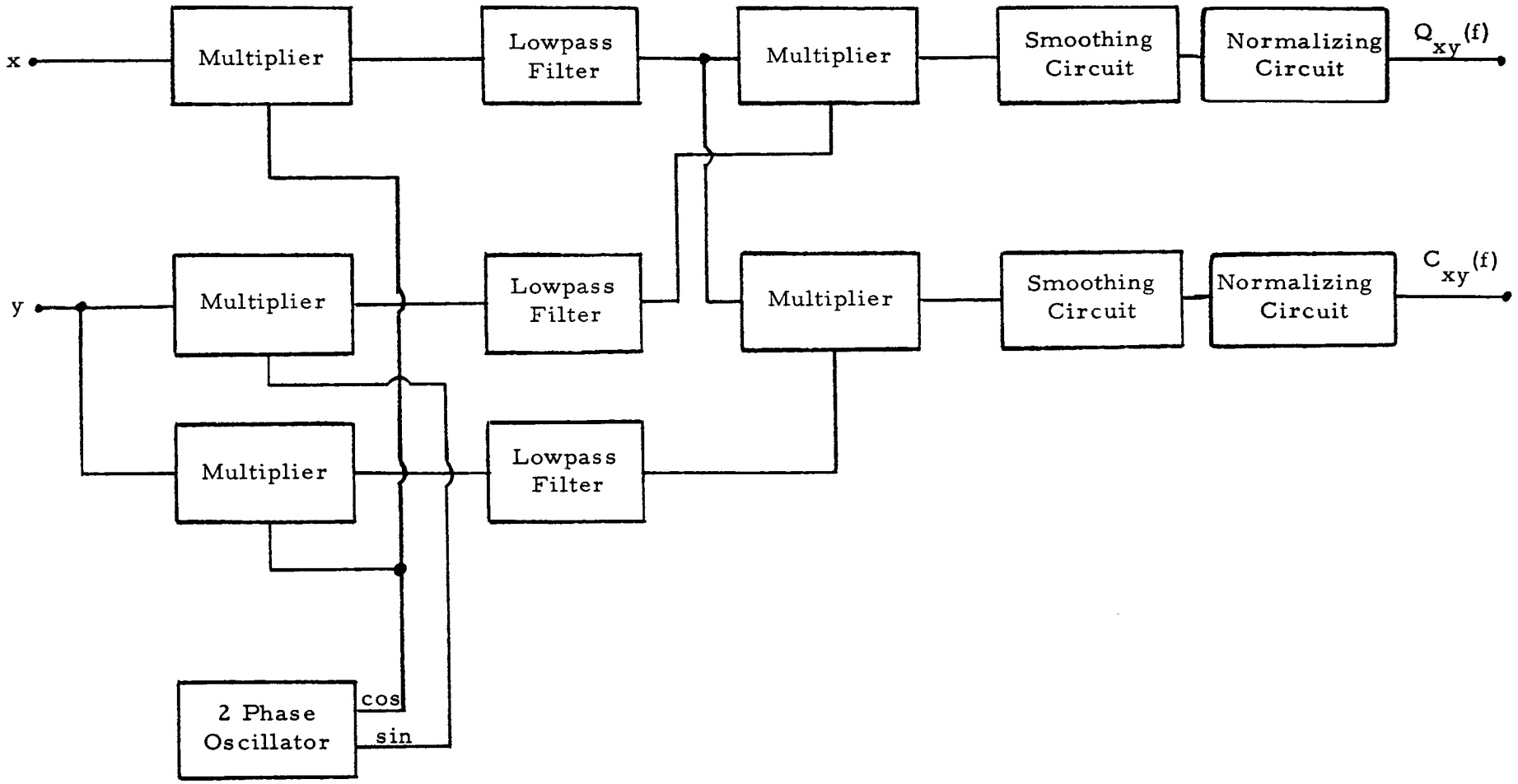


Figure 3-15. Zero IF Frequency Cross Spectral Density Analyzer

$$G_{xy}(f) = \frac{1}{T}\left[\int_{-T}^{T} x(t)\left\{\cos 2\pi ft - j \sin 2\pi ft\right\} dt\right]\left[\int_{-T}^{T} y(t)\left\{\cos 2\pi ft + j \sin 2\pi ft\right\} dt\right] \tag{3.22a}$$

$$\begin{aligned} G_{xy}(f) = \frac{1}{T}\Bigg[&\left(\int_{-T}^{T} x(t) \cos 2\pi ft\, dt\right)\left(\int_{-T}^{T} y(t) \cos 2\pi ft\, dt\right) \\ &+ \left(\int_{-T}^{T} x(t) \sin 2\pi ft\, dt\right)\left(\int_{-T}^{T} y(t) \sin 2\pi ft\, dt\right)\Bigg] \\ &- j\left(\int_{-T}^{T} x(t) \sin 2\pi ft\, dt\right)\left(\int_{-T}^{T} y(t) \cos 2\pi ft\, dt\right) \\ &- \left(\int_{-T}^{T} x(t) \cos 2\pi ft\, dt\right)\left(\int_{-T}^{T} y(t) \sin 2\pi ft\, dt\right)\Bigg] \end{aligned} \tag{3.22b}$$

The real portion of Eq. (3.22b) is the cospectrum and the imaginary portion is the quadspectrum. Thus one can compute the co- and quadspectra by multiplying x(t) and y(t) by sine and cosine terms and performing the operations shown in Eq. (3.22b).

A difficulty encountered with directly implementing Eq. (3.22b) is that the filtering is accomplished by the integration process; hence, the effective bandwidth of the analyzer is totally dependent upon the integration

time. A modification that is used in commercial analyzers consists of replacing all of the integration operations with lowpass filters. (These filters are made as nearly identical as possible.) Then the raw cospectra and the raw quadspectra are passed through an additional stage of lowpass filtering before display. A block diagram of a typical cross-spectral density analyzer based on direct Fourier transformation is shown in Figure 3-16.

3.3 CONCLUSIONS AND RECOMMENDATIONS

The performance of analog correlators in general leaves a lot to be desired, particularly when compared to the hybrid correlators discussed in Section 5. Probably the major problem area is that of accurately implementing the time delay. With drum delay correlators one has the problems of maintaining very strict dimensional tolerances in a complex mechanical mechanism, short record lengths for high frequency data on a reasonable drum diameter, and frequently, excessive drum bearing wear. Magnetic tape delay correlators are limited by nonuniform tape velocity. The magnetic correlator described is limited in upper frequency response and the optical correlators are presently operationally inconvenient.

The performance of analog cross-spectral density analyzers, on the other hand, is quite good. Of the types described, those employing zero IF filtering techniques are recommended as the best presently available because this technique permits extremely sharp drift-free filtering of the data. It should be noted however, that since these analyzers consist of a single pair of filters that are scanned across the frequency range of analysis, an appreciable amount of analysis time is required.

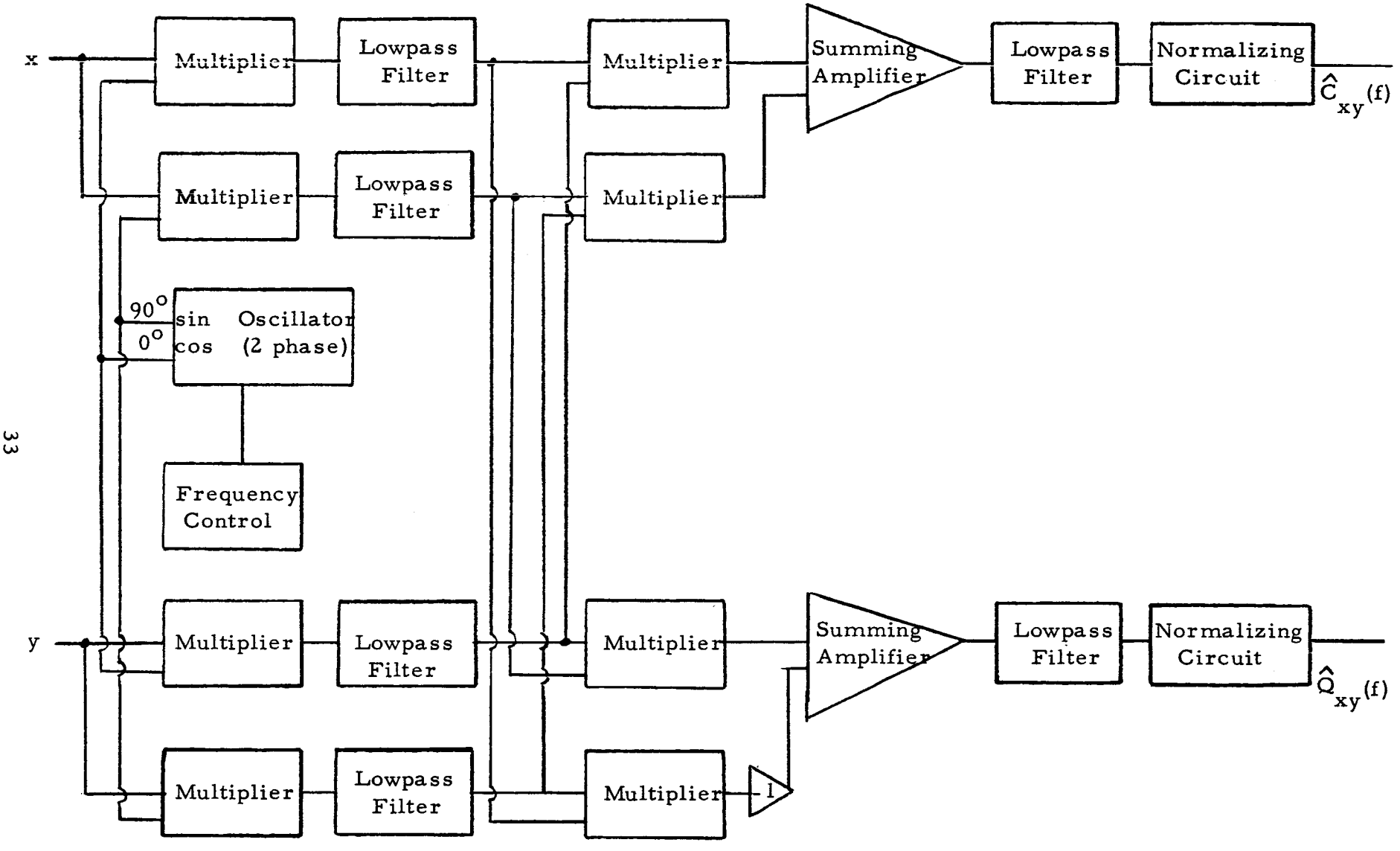


Figure 3-16. Cross-Spectral Density Analyzer Using Direct Fourier Transform Techniques

4. DIGITAL DATA REDUCTION TECHNIQUES

4.1 INTRODUCTION

The object of this section is to compile and describe many of the digital techniques that have been developed to perform the analyses necessary for time series data. Emphasis has been placed on techniques that minimize the required computational time. The need for computational efficiency arises from the fact that correlation and spectral density calculations (to be defined more precisely in following sections) require, in general, large amounts of computation. The primary computational loop that is involved is a multiply-add sequence which often must be executed on the order of 10^7 times. This means, for example, that on a modern high-speed large-scale digital computer where a fixed point multiply-add loop could be executed in, say, 20 microseconds, a total of 200 seconds, or 3 minutes and 20 seconds would be required for just the execution of this loop. This would be the computational time required for just one single data record and, of course, in many physical applications the number of data records to be analyzed both individually and jointly can run into fairly large numbers. Hence, there is a real need to give attention to this correlation computation loop in order to make it less time consuming and more efficient.

Since this discussion will be primarily concerned with the discrete case given by Eq. (2.3), the notation will be simplified for later use. That is, it will be understood that the interval $\Delta\tau$ separates two consecutive points of the computed correlation function, and hence all reference to $\Delta\tau$ or Δt will be dropped. Equation (2.3) then becomes

$$\hat{R}_{xy}(r) = \frac{1}{N-r} \sum_{i=1}^{N-r} x_i \, y_{i+r} \quad , \quad r = 0, 1, 2, \ldots, m \qquad (4.1)$$

where $\hat{R}_{xy}(r) = \hat{R}_{xy}(r\Delta\tau)$

$$x_i = x(i\Delta t)$$

$$\Delta\tau = \Delta t$$

The most widely used digital computational method for spectra is based on taking the discrete Fourier transform of Eq. (4.1). In practice, the Fourier transform takes a negligible amount of computational time as compared to the correlation computations. The potentially time consuming cosines and sines may be either pre-calculated or generated by an efficient recursive procedure which takes very little time. Thus, the over-all computational time is small in comparison with the correlation function computations.

Recent developments and improved methods of digital or numerical filtering have led to procedures whereby the power spectral density function is computed directly by simulating the analog method. That is, the process $x(t)$ is filtered in a narrow frequency band B_a by the appropriate digital filter and then squared, averaged and normalized to the analysis bandwidth B_a. In the succeeding subsections, methods for both efficient correlation computations and efficient filtering procedures will be discussed. A more precise definition and more detailed explanation of digital filtering will be deferred until Section 4.3.

It should be noted at this time that filtering procedures amount to taking a sum of cross products between a sequence of predetermined weights a_i and the process x_i so that both correlation and filtering computations reduce to multiply-add sequences. Hence, it is this operation that is of fundamental importance for either method.

Storage of the data represents a serious problem for any digital analysis of data containing high frequencies or requiring small time delay resolution. Very high digitizing rates are necessary to meet either of the above two requirements. If the data contains low frequencies, the sample record length must be as long or longer than three or four periods of the lowest frequency

component to prevent excessive bias errors from these frequencies. With high digitizing rates, multiple channels and moderately long sample record lengths, core storage capacity will be greatly exceeded and in many cases it is easy to exceed the total storage capacity of digital tapes.

4.2 CORRELATION FUNCTION COMPUTATIONS

In practice, two different types of time series data arise which are to be processed by a digital computer. Some processes are discrete by their very nature, say, for example, the daily closing prices of a given stock. On the other hand, some data arises naturally as a continuous record such as the continuously recorded output of an accelerometer located on the skin of a missile structure which is intended to give a measure of the vibration at that point. The continuous record is digitized by an analog to digital conversion procedure for digital analysis.

In either case, an important observation to make is that only a finite number of binary digits (bits) in a binary digital computer are required to represent any given individual data point. In actual practice, many analog to digital converters present their output to a precision of 8-10 bits (including the sign bit). For almost all applications, the recording instruments are no more accurate than this, (one part in 256 to 1024) and hence, this is a fine enough quantization.

As it turns out, for many time series parameters of interest, much coarser quantization is often acceptable. In the extreme case, one can quantize to a single bit. That is, if the value of a signal is larger than or equal to zero, set the bit equal to zero; if the value of the signal is less than zero, set the bit equal to one. In simpler terms, the sign bit of the data point is the only information retained. Applications of this concept will be discussed in Section 4.2.4. In mathematical terms, if $x(t)$ is the original signal, then the one bit quantized (hard clipped) signal $y(t)$ is defined by the relation

$$y(t) = \operatorname{sgn} x(t) = \begin{cases} 1 \;, & x(t) \geq 0 \\ -1 \;, & x(t) < 0 \end{cases} \tag{4.2}$$

A very important observation which is made is that when 8 to 10 bits or less are employed for the level of quantization, the number of different values that a data point can take on is not too great, namely $2^8 = 256$ to $2^{10} = 1024$ different values. By proper development of this concept, certain time saving procedures can be implemented and are discussed in Sections 4.2.2, 4.2.3, and 4.2.5.

In the following subsections, different procedures for computing the correlation function are discussed. First, the basic direct method of implementing Eq. (4.1) is described along with an indication of the number of operations necessary for its implementation. The other methods are:

(a) one which actually expands the summation and collects like terms so that common factors may be factored out in order to trade additions for multiplications;

(b) a method which involves relatively coarse quantization in order that advantage may be taken of one of the special convert instructions of the IBM 7090 class machines which in effect perform the table look-up;

(c) so-called hard clipping methods which are based on the one-bit quantization;

(d) a method which takes advantage of the fact that a cross product can be expanded as a linear combination of squares which can then be efficiently obtained by table look-up procedures; and finally,

(e) a direct Fourier transformation of the original time history to obtain the power spectrum and then the correlation function.

These procedures will now be described.

4.2.1 Basic Correlation and Spectrum Computational Method

The fundamental correlation computational procedures consist of implementing Eq. (4.1) directly. Then, if the spectral density function is

desired, as is usually the case, the discrete Fourier transform is taken. The formula for $\hat{R}_{xy}(r)$ is repeated here for convenience.

$$\hat{R}_{xy}(r) = \frac{1}{N-r} \sum_{i=1}^{N-r} x_i \, y_{i+r} \quad , \quad r = 0, 1, \ldots, m$$

In practice, since $R_{xy}(\tau) = R_{yx}(-\tau)$, two parts of the cross-correlation function will be computed for positive time delays. That is,

$$\hat{R}_{xy}(r) = \frac{1}{N-r} \sum_{i=1}^{N-r} x_i \, y_{i+r} \quad , \quad r = 0, 1, \ldots, m \qquad (4.3a)$$

$$\hat{R}_{yx}(r) = \frac{1}{N-r} \sum_{i=1}^{N} y_i \, x_{i+r} \quad , \quad r = 0, 1, \ldots, m \qquad (4.3b)$$

The spectrum will be obtained from the correlation function but since only positive frequencies are of physical interest, a physically realizable spectrum, $G_{xy}(f)$, is defined as two times $S_{xy}(f)$ for positive frequencies only. That is,

$$G_{xy}(f) = 2S_{xy}(f) = 2 \int_{-\infty}^{\infty} R_{xy}(\tau) \, e^{-j 2\pi f \tau} \, d\tau \quad ; \quad 0 \leq f < \infty \qquad (4.4)$$

$$= 2 \int_{-\infty}^{\infty} [R_{xy}(\tau) \cos 2\pi f\tau - j R_{xy}(\tau) \sin 2\pi f\tau] \, d\tau \quad ; 0 \leq f < \infty$$

The discrete version will be computed in terms of the real and imaginary parts, the co-spectra and quad-spectra, as follows. In the subsequent

discussion, the hat will be dropped from $R_{xy}(r)$ to simplify notation.

$$C_{xy}(k) = 2\Delta\tau\left[A_{xy}(0) + 2\sum_{r=1}^{m-1} A_{xy}(r)\, D(r) \cos\frac{\pi rk}{m} + (-1)^k A_{xy}(m)\right] \qquad (4.5)$$

$$Q_{xy}(k) = 4\Delta\tau \sum_{r=1}^{m-1} B_{xy}(r)\, D(r) \sin\frac{\pi rk}{m} \quad , \quad k = 0, 1, \ldots, m \qquad (4.6)$$

where for notational convenience $C_{xy}(k\Delta f) = C_{xy}(k)$ and $D(r)$ will be defined shortly. The "even" and "odd" parts of $R_{xy}(\tau)$ are computationally convenient and are defined as

$$A_{xy}(r) = \frac{1}{2}\left[R_{xy}(r) + R_{yx}(r)\right] \qquad (4.7)$$

$$r = 0, 1, 2, \ldots, m$$

$$B_{xy}(r) = \frac{1}{2}\left[R_{xy}(r) - R_{yx}(r)\right] \qquad (4.8)$$

where $R_{yx}(r) = R_{xy}(-r)$. The computational forms given by Eqs. (4.5) and (4.6) are trapezoidal rule discrete integration forms and are standardly used. An additional multiplicative factor of two appears in Eqs. (4.5) and (4.6) since the Fourier transform integrations are being performed over positive time only.

The function $D(r)$ is termed a "lag window" and is necessary due to certain statistical considerations. This function results in good statistical estimates being obtained within a small bandwidth B_a centered at frequency f rather than a poor statistical estimate of nearly zero bandwidth at frequency f. This function may be related to the gain factor of the filter used in the direct filtering method of spectral computations. For further discussion, see Reference 10. Simplifications of Eqs. (4.1), (4.5), (4.6) and (4.7) result

in the autocorrelation case when $y(t) = x(t)$. Then, since $R_{xx}(-r) = R_{xx}(r)$, the computations for negative values of r need not be performed. Also, $A_{xx}(r) = R_{xx}(r)$ and $B_{xy}(r) = 0$ so the imaginary part of $G_{xx}(f)$ is zero. Hence, $G_{xx}(f)$ is a real number.

As mentioned before, the primary computational time is spent in the evaluation of the correlation function. The Fourier transform, including the so-called lag window, takes a nearly negligible amount of computational time if the following recursive scheme is used for the evaluation of the sine/cosine values necessary in the cross-spectral density computations.

$$\begin{bmatrix} \cos k(i+1)\,\pi/m \\ \sin k(i+1)\,\pi/m \end{bmatrix} = \begin{bmatrix} \cos \pi/m & -\sin \pi/m \\ \sin \pi/m & \cos \pi/m \end{bmatrix} \begin{bmatrix} \cos k\pi i/m \\ \sin k\pi i/m \end{bmatrix} \qquad (4.9)$$

$$i = 0, 1, \ldots, m$$

where k is the frequency value index.

For the special case of a power spectral density function, only cosines are needed which may be obtained from the recursion

$$\cos \frac{k\pi(i+1)}{m} = \left(2 \cos \frac{k\pi i}{m}\right)\left(\cos \frac{k\pi}{m}\right) - \cos \frac{k\pi(i-1)}{m} \qquad (4.10)$$

$$i = 1, \ldots, m$$

These recursion formulas save significant amounts of time as compared to going to a sine/cosine function subroutine for each necessary value. However, special purpose sine/cosine subroutines to give only the necessary accuracy might be as efficient or more so. The number of multiply-add's required for the computation of $R_{xy}(r)$ is roughly $(m+1)\,(N - \frac{m}{2}) \approx mN - \frac{m^2}{2}$. This essentially defines the required computation time. For the IBM 7094, the

innermost multiply-add loop in floating point arithmetic takes roughly 12 machine cycles at 1.4 μsec per cycle. Thus, for example, a typical program operating on N = 10,000 data points for m = 1000 lags, the computation time could be estimated to be

$$T_a = nc\,T_1$$
$$= [mN - \frac{m}{2}^2]\,[12]\,[1.4 \times 10^{-6}]$$
$$= [10^4 \times 10^3 - \frac{10^6}{2}]\,[12]\,[1.4 \times 10^{-6}] \approx 160 \text{ sec.}$$

where:

T_a = the total computation time

n = the number of multiply-add operations

c = the number of machine cycles per multiply-add operation

T_i = the time for one machine cycle

If the sampling interval was $\Delta t = 0.1$ millisecond, then the parameters in the example correspond to a 10 Hz analysis bandwidth performed in the frequency band 0 to 5000 Hz. See Reference 11 for a further discussion.

4.2.2 Factoring of Common Terms

A method employed at MIT in connection with seismic data analysis is described in Reference 12. This method utilizes a data scan to factor out common terms which appear in the summation for the evaluation of the correlation function. This procedure can only be efficiently accomplished if a relatively small number data quantization levels, as compared to the total number of data points, are possible so that relatively large numbers of common factors exist on the average. For example, suppose one has obtained a sequence of data as indicated below.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	3	-2	1	4	2	2	3	4	1	0	1	-1	3	2	-4	0	3	4	1	-2

One can expand Eq. (2.1) in the form

$$
\begin{aligned}
R_{xy}(r) = \; & 1 \cdot [x_{3+r} + x_{9+r} + x_{11+r} + x_{19+r} - x_{12+r}] \\
& +2 \cdot [x_{5+r} + x_{6+r} + x_{14+r} - (x_{2+r} + x_{20+r})] \\
& +3 \cdot [x_{1+r} + x_{7+r} + x_{13+r} + x_{17+r}] \\
& +4 \cdot [x_{4+r} + x_{8+r} + x_{18+r} - x_{15+r}]
\end{aligned}
\qquad (4.11)
$$

Note that when Eq. (2.1) has been rewritten in this form, only 4 instead of 20 multiplications are necessary. Due to the reduction in the number of multiplies, the evaluation of Eq. (4.11) can be performed considerably faster than Eq. (2.1). This is where the potential advantage lies.

Note that Eq. (4.11) is a function of the specific data being used. Therefore, the particular record of data which is being analyzed must be examined in order to generate the computational form given by Eq. (4.11). A complicated program writing program is required; that is, a program which will generate a computational program of the form of Eq. (4.11) for each given input data record.

A correlation function program involved in this procedure was written for the data analysis procedures described in Reference 12. Assuming a quantization of 100 levels, (between 6 and 7 bits) empirical speed advantage factors were determined which are a function of the length of the data, as indicated in the table below.

Table 4.1 Speed Advantage (SA) Ratio of Computing Time of Normal Method to Factoring Method

N	500	1000	2000	5000	10,000
SA	3.6	5.7	8.1	10.9	12.4

Hence, for the particular data involved in this application, a considerable savings in time resulted from this highly specialized computational method. Considerable programming effort is required to develop the so-called program writing program as opposed to merely coding the direct computational method which is a disadvantage. This factor would have to be balanced against the time savings which would be realized by this special method.

4.2.3 Eight Level Quantization Method

Personnel at the Jet Propulsion Laboratory have programmed a method based on a 4-bit quantization level. That is, plus or minus 8 levels of quantization are allowed which corresponds to roughly plus or minus one decimal digit. The advantage of this coarse quantization is that if a few enough number of levels is used, then table look-ups rather than multiply operations can be employed in order to determine the products between a pair of data values. Of course, when the number of potential products between two arbitrary data values is too large, then it is likely that core storage is not large enough to contain the necessary table and this procedure cannot be used. Many of the fast correlation computation methods are based on being able to employ a table look-up procedure instead of direct multiplication. As will be shown in the next subsection, data can be quantized all the way to only one bit and still reasonable results can be obtained by a final functional adjustment of the results.

The Jet Propulsion Laboratory Computation Center utilizes an IBM 7094 computer. A special instruction on the IBM 7094 provides an automatic table look-up which may be utilized for performing the multiplications automatically. The particular instruction used is one of the convert instructions (CAQ). By proper construction of the multiplication table of possible multiplication results between two data points and by proper placement of the two sets of data to be multiplied, one can utilize one of these convert instructions to perform the

table look-up to obtain the desired products between two sets of data. Similar procedures are used for the convert instructions in order to convert from binary to binary coded decimal formats. The data must be pre-processed in order to properly pack several data points into one computer word. The convert instruction essentially adds in sequence the six, six-bit characters in a computer word to a predetermined value and then accumulates the contents of this location in the accumulator. Hence, six pairs of data points can be multiplied in one instruction which is accomplished much faster than regular multiply operations.

The results of this procedure, based both on experimental and analytical results, indicate that very little precision is lost due to this coarse quantization. This is to be intuitively expected since the correlation function is an averaged result and results with more precision than the original data are often not of any particular interest. But when one adds up many numbers of a given accuracy, one, in effect, gains significant digits of precision and the mean values resulting contain more accuracy than the original data. This additional accuracy is, in general, of no interest.

One disadvantage encountered in employing such a procedure, and, as a matter of fact, the reason many engineering personnel do not like the results of the 4-bit quantization method, is that so-called dynamic range is lost. That is, one has a fundamental noise floor, so to speak, which increases in proportion to the minimum level of quantization.

The JPL 8-level method has an inner computational loop requiring 24 machine cycles to obtain the sum of six products. As an estimate, an additional two machine cycles might be required on the average to prepare the data in the necessary six data point per computer word format. The total is then 26/6 cycles per multiply add. For the previous example of N = 10,000 and m = 1,000 the timing is approximately

$$T = \frac{160}{12} \cdot \frac{26}{6} = 57.8 \text{ sec}$$

4.2.4 One-Bit Quantization Methods (Extreme Clipping Methods)

A relation between the correlation function of the so-called extreme clipped signal [see Eq. (4.2)] and an original Gaussian zero mean process from which it is obtained gives the normalized correlation function (correlation coefficient) $\rho_{xx}(\tau)$ in terms of the correlation coefficient function of the clipped signal $\rho_{yy}(\tau)$ by the formula

$$\rho_{xx}(\tau) = \sin\left[\frac{\pi}{2}\,\rho_{yy}(\tau)\right] \tag{4.12}$$

where $\rho_{xx}(\tau)$ and $\rho_{yy}(\tau)$ are defined as

$$\rho_{xx}(\tau) = \frac{R_{xx}(\tau)}{R_{xx}(0)} \quad ; \quad \rho_{yy}(\tau) = \frac{R_{yy}(\tau)}{R_{yy}(0)} \tag{4.13}$$

See page 34 of Reference [13] for a derivation of this relation.

Clearly, when the one-bit quantization is performed, the multiplications which need be accomplished become trivial. Various ways exist for taking advantage of this trivial multiplication. One method which is employed by JPL is to again use one of the convert instructions of the IBM 7094. In this case the data is packed, 36 points per word, by an initial examination of the input record. Then by proper generation of the table look-up procedure and by employing the proper table, one can, in effect, look up 36 products with one instruction by use of a convert operation. Six sets of products at a time are obtained since the convert instruction operates on 6-bit character sets.

Other procedures than the use of the convert instruction might prove to be just as efficient for evaluating a sum of one-bit products. One might easily set up simple counters where after a point, say x_i is selected, then

all the products of x_i with each data point y_{j+r}, $j = 1, \ldots, N-r$, could be determined by simple examination of the data. This problem is not one that can be solved in any general manner, because it depends on the characteristics of the particular computer being used for the problem at hand.

There are certain problems connected with the use of the clipping methods for correlation computations. One problem is that increased record length requirements are traded for computational speed assuming one wants to maintain a constant statistical accuracy. It is shown in Reference 13, page 37, that the variance of an autocorrelation function estimate is increased by a maximum factor of roughly $\pi^2/4 \approx 2.5$. This variance is proportional to T, so if the variance is to be maintained constant, one requires record lengths which are two and a half times as long as those needed if all the information available in the data is to be used. As is further noted in Reference 13, roughly the same bound holds approximately true for power spectral density estimates. Although these are not precise theoretical relationships, they provide one with guidelines for evaluating the tradeoff between accuracy and computational time when employing clipping type correlation function computations.

An additional comment concerning the coarsely quantized correlation function computation comes from the University of Arizona. In the Electrical Engineering Department there, some experiments have been made testing the variability of methods employing quantitized data. It has been found that rarely over a 150% increase in the variability of autocorrelation function estimates was experienced. Although the theoretical bound is a maximum increase of 250%, it seems to be much less than this in practice. However, there are apparently no completely acceptable analytical results available on the variability questions at this time.

Note that certain assumptions are involved in the use of the hard clipping method:

(a) Either a Gaussian distribution is assumed, or

(b) a sinusoid, or

(c) a sinusoid plus Gaussian noise.

Equation (4. 12) applies to data with these characteristics. Note that any periodic data with the same period (i. e. , the same zero crossings) would give the same clipped correlation function as that of a sine wave whose clipped correlation function is a triangular wave prior to applying Eq. (4. 12). In addition, only normalized correlation functions and spectra are obtained. Thus, the mean square value must be calculated separately for use as a scale factor. Note that nonlinear effects which distort probability density functions to non-Gaussian shapes would completely be masked when such a computing method is used.

An alternative method employing the clipping idea to correlation function computations is that of clipping only one of the components of the product being computed. For example, if the cross-correlation function is to be computed between x(t) and y(t) then clip only x(t). In this case one again trades multiplications for adds in that one merely adds up the sequence y_i when adjusted for the proper sign of the point x_i which is multiplying it. In this case the correction factor can be shown to be

$$R_{xy}(\tau) = \sqrt{\frac{\pi}{2}}\, R_{yy}(0)\, R_{x\ \mathrm{sgn}\ y}(\tau) \qquad (4.14)$$

where $R_{x\ \mathrm{sgn}\ y}(\tau)$ denotes the correlation function of x(t) with the hardclipped version of y(t).

This procedure has been tested in the Vibration Data Systems Section of Marshall Space Flight Center. Encouraging empirical results have been obtained as long as the records are reasonably Gaussian and as long as the analyses are performed over reasonably long record lengths.

Relative time estimates for this method are given in Reference 14. For the IBM 7044 it is indicated that the computing time can be reduced from 268 to 41 seconds which is a ratio of about 6 to 1. If this ratio is assumed to hold for the IBM 7094 and is applied to the $N = 10,000$, $m = 1000$ example, the final computing time would be

$$T = \frac{160}{6} = 26.7 \text{ sec}$$

Presumably, the full clipping method would take proportionately less time. Although specific coding is not available, one could expect to perform 36 multiply-add operations in roughly the same 24 machine cycle loop as referred to in Section 4.2.3. However, the data preparation would likely require more time to prepare in the proper format. As a conservative estimate, assume that an average of 24 additional machine cycles are necessary or a total of 60 cycles per 36 multiply-adds. The total time for the $N = 10,000$, $m = 1000$ example would then be

$$T = \frac{160}{12} \cdot \frac{60}{36} = 22.2 \text{ sec}$$

4.2.5 Sum of Squares Method

An additional technique which has been suggested (see Reference 15) takes advantage of the finite quantization of the data and expresses the product as a sum of squares. The fundamental relation which is employed here expresses the fact that a cross product may be expressed as a linear combination of squares by the following relation:

$$xy = \frac{1}{2}\left[(x+y)^2 - x^2 - y^2 \right] \qquad (4.15)$$

This is very similar to the fairly well known "quarter-square" method which has been used in construction of analog multipliers. (See Eq. 3.4 and subsequent discussion,) This equation may be modified into a form involving the correlation function immediately.

$$R_{xy}(r) = \frac{1}{2N}\left[\sum_{i=1}^{N-r}(x_i + y_{i+r})^2 - \sum_{i=1}^{N-r} x_i^2 - \sum_{i=1}^{N-r} y_{i+r}^2\right] \qquad (4.16)$$

The key point in using such a relation is the fact that tables of squares take up much less storage space than tables of cross products for the same number of data points. For example, say 10-bit quantization is being used so that a table of size 2^{20} cells would be required if all possible cross products were to be stored. The same table of squares of the sum of the data points would only take up 2^{11} cells. 2^{11} might be a very reasonable table size, whereas 2^{20} is not.

The second and third sums of squares on the right hand side of Eq. (4.16) are evaluated by simple recursion relations. First, one computes the sum of squares of all the data. This requires only one pass through the data. Then the partial sums of squares are obtained from

$$\sum_{i=1}^{N-r} x_i^2 = \sum_{i=1}^{N-r+1} x_i^2 - x_{N-r+1}^2 \quad , \; r = 1, 2, \ldots, m$$

$$\sum_{i=1}^{N-r} y_{1+r}^2 = \sum_{i=1}^{N-r+1} y_{i+r-1}^2 - y_r^2 \quad , \; r = 1, 2, \ldots, m$$

Since only one pass through the data is required initially plus one subtraction operation for each correlation function point, the amount of time required is negligible compared with the over-all correlation function evaluation.

The other portion of Eq. (4. 16) requires a more complicated method for evaluation. One possible method of programming for the IBM 7090 is given in Reference 15. This method requires one pass through the data to set it up in a special tabular format. Then with the aid of a table of squares, the quantity $\sum_{i=1}^{N-r} (x_i + y_{i+r})^2$ can be generated in a computational loop requiring six machine cycles. If one allows one extra machine cycle to account for the table set up, etc., which must be done, the computational time estimate for the example of Section 4. 2. 1 ($N = 10^4$, $m = 10^3$) would be about

$$T = \frac{7}{12} \ 160 \text{ sec} \approx 93.3 \text{ sec}$$

whereas the fixed point arithmetic loop for the basic way requires 11 cycles and the time is

$$T = \frac{11}{12} \ 160 = 146.7 \text{ sec}$$

There are certain additional considerations with this procedure. First, it is assumed the data is in fixed point and second, when the amount of data is too large to fit into core memory simultaneously, some of the advantage might disappear.

4. 2. 6 Direct Computation of Spectrum from Fourier Transform

The unsmoothed or raw spectral density function (the periodogram) can be computed directly from the Fourier transform of the original time history. The power spectral density function can be defined as the variance of the transformed process. The cross-spectrum is the covariance between

two transformed variables. In equation form these statements are

$$\tilde{G}_{xx}(f) = \mathrm{Var}[X(f)] = E[\,|X(f)|^2] \tag{4.17}$$

$$\tilde{G}_{xy}(f) = \mathrm{Cov}[X(f), Y(f)] = E[X(f)\, Y^*(f)] \tag{4.18}$$

The wavy line is to indicate the unsmoothed spectrum, the asterisk denotes complex conjugate. For discrete time records, the averaging (expectation) is done over time. The computational forms of Eqs. (4.17) and (4.18) are

$$\tilde{G}_{xx}(k) = \Delta t\left\{\left[\frac{x_0}{2} + \sum_{i=1}^{N-1} x_i \cos\frac{2\pi ki}{N}\right]^2 + \left[\sum_{i=1}^{N-1} x_i \sin\frac{2\pi ki}{N}\right]^2\right\} \tag{4.19}$$

$$k = 0, 1, 2, \ldots, N-1$$

$$\tilde{G}_{xy}(k) = \Delta t\left\{\left[\frac{x_0}{2} + \sum_{i=1}^{N-1} x_i \cos\frac{2\pi ki}{N}\right] - j\left[\sum_{i=1}^{N-1} x_i \sin\frac{2\pi ki}{N}\right]\right\} \tag{4.20}$$

$$\left\{\left[\frac{y_0}{2} + \sum_{i=1}^{N-1} y_i \cos\frac{2\pi ki}{N}\right] + j\left[\sum_{i=1}^{N-1} y_i \sin\frac{2\pi ki}{N}\right]\right\}$$

$$k = 0, 1, \ldots, N-1$$

To obtain final smoothed estimates equivalent to those obtained from Eqs. (4.5) and (4.6), a smoothing (convolution) operation must be performed. That is, one evaluates the equations

$$G_{xx}(k) = a_0 G_{xx}(k) + \sum_{i=1}^{\ell} a_i [\tilde{G}_{xx}(k+i) + \tilde{G}_{xx}(k-i)]$$

$$k = 0, \frac{N}{m}, \frac{2N}{m}, \ldots, \frac{m-1}{m} N, N$$

where the a_i, $i = 0, 1, \ldots, \ell$ are the discrete values of the Fourier transform of some particular lag window $D(r)$.

This method is usually not considered because of the large computational requirements for direct evaluation of Eqs. (4.19) and (4.20). For example, the power spectrum requires about 2N multiply-add operations for each of N raw spectrum values. This is a total of $2N^2$ operations for all N points. The generation of the necessary sines and cosines would require additional time if it is assumed that the values are not already available in storage. Equation (4.9) would be used for sine/cosine generation. Thus, $2N^2$ operations plus additional smoothing are to be compared with the Nm multiply-add operations necessary when computing the correlation function plus a Fourier transform of the correlation function. Thus, for the case of N = 10,000 points, the computational time is estimated to be T = 1600 seconds.

One gains the maximum resolution or, equivalently, narrowest possible bandwidth ($B_a = 1/T$) for the additional computing time. However, due to the large statistical variability associated with narrow bandwidths, it would be assumed that wider bandwidth estimates are generally wanted. The smoothing indicated by Eq. (4.21) reduces the statistical variability while widening the bandwidth.

4.2.7 Cooley-Tukey Variation of the Direct Fourier Transform Method

Special methods are available under certain conditions to speed the computation of a Fourier transform. A method is described in Reference 16 which reduces $2N^2$ to roughly $2N\log_2 N$ complex multiply-add operations where N is an even power of 2, that is,

$$N = 2^p, \text{ p an integer} \qquad (4.22)$$

Since the multiplication and addition of two complex variables requires about four times the work of a real variable multiply-add, the number of operations becomes $8N\log_2 4N$ in terms of real numbers.

The algorithm given in Reference 16 applies generally in complex numbers, but no particular simplification exists for spectrum computations based on only real numbers. The computational method is a recursion and intermediate values become complex numbers which obviates significant simplifications being made. Let the Fourier transform be denoted by $X(k) = X(k\Delta f)$ and defined by the equation

$$X(k) = \Delta t \sum_{i=1}^{N-1} x_i \, e^{-j\frac{2\pi ik}{N}} \quad , \quad k = 0, 1, \ldots, N-1 \tag{4.23}$$

where N is a power of two defined by Eq. (4.22). The input data x_i and the results X(k) can be considered to be stored in the memory of a binary digital computer. Then the indices k and i can be expressed as binary expansions.

$$k = k_{p-1}\, 2^{p-1} + \ldots + k_1 2 + k_0$$

$$i = i_{p-1}\, 2^{p-1} + \ldots + i_1 2 + i_0 \tag{4.24}$$

$$k_\nu = 0, 1 \quad ; \quad i_\nu = 0, 1$$

The numbers X(k) and x_i are considered as being stored in binary core memories with location addresses given in binary form beginning at zero as indicated in Figure 4-1, which is an example for N = 16, p = 4.

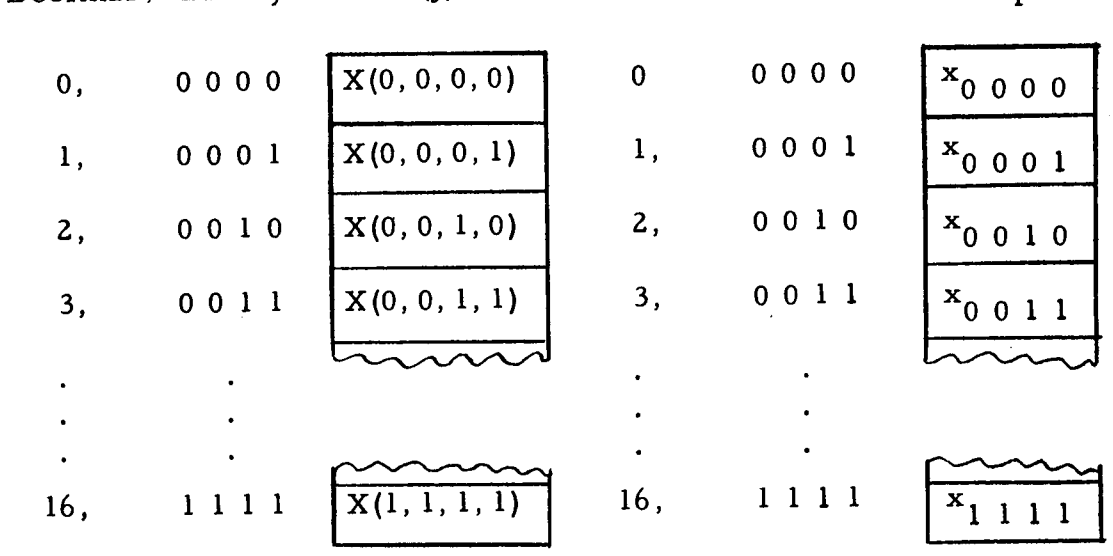


Figure 4.-1 Binary Storage Locations for Fourier Transform

Now Eq. (4.23) can be written

$$X(k_{p-1},\ldots,k_0) = \sum_{k_0}\sum_{k_1}\cdots\sum_{k_{p-1}} x_{i_{p-1},\ldots,i_0}\, e^{-\frac{j2\pi k}{N}(i_{p-1}\cdot 2^{p-1}+\ldots+i_0)} \tag{4.25}$$

$$k_\nu = 0, 1$$

Each summation consists of only two terms since one is adding over the binary components of the number. These summations can be performed sequentially by making use of the relation

$$e^{-j\frac{2\pi}{2^p}k\cdot i_{p-1}2^{p-1}} = e^{-j2\pi\frac{k_{p-1}2^{p-1}i_{p-1}}{2}}\, e^{-j2\pi\frac{k_{p-2}2^{p-2}i_{p-1}}{2}}\cdots e^{-j\frac{2\pi}{2^p}k_0 i_{p-1}2^{p-1}}$$

$$= (1\cdot 1\cdot\ldots\cdot 1)\, e^{-j\frac{2\pi}{N}k_0 i_{p-1}2^{p-1}} \tag{4.26}$$

Since $i_{p-2}, i_{p-3}, \ldots, i_0$ have been removed from the exponential, one can evaluate the two term sum

$$A_1(k_0, i_{p-2}, \ldots, i_0) = \sum_{i_{p-1}} x_{i_{p-1}, i_{p-2}, \ldots, i_0} e^{-j\frac{2\pi}{N} k_0 i_{p-1} 2^{p-1}} \tag{4.27}$$

$$k_0 = 0, 1; \quad i_\nu = 0, 1$$

Note that A_1 consists of a set of N numbers each depending on two of the original data points. In the example of Figure 1, x_{0000} and x_{1000} would be involved in both $A_1(0, 0, 0, 0)$ and $A_1(1, 0, 0, 0)$. The general recursion is based on the following generalization of Eq. (4.26).

$$e^{-j\frac{2\pi}{N} k \cdot i_{p-\ell} 2^{p-\ell}} = e^{-j\frac{2\pi}{N}\left(k_{\ell-1}2^{\ell-1}+\ldots+k_0\right) i_{p-\ell} 2^{p-\ell}} \tag{4.28}$$

Then successive sums are evaluated according to the equation

$$A_\ell(k_0, \ldots, k_{\ell-1}, k_{p-\ell-1}, \ldots, i_0) \tag{4.29}$$

$$= \sum_{i_{p-\ell}} A_{\ell-1}(k_0, \ldots, k_{\ell-2}, i_{p-\ell}, \ldots, i_0)\, e^{-j\frac{2\pi}{N}(k_{\ell-1}2^{\ell-1}+\ldots+k_0) i_{p-\ell} 2^{p-\ell}}$$

In its simplest computational form, Eq. (4.29) becomes

$$A_\ell(k_0,\ldots,k_{\ell-1},i_{p-\ell-1},\ldots,i_0) = A_{\ell-1}(k_0,\ldots,k_{\ell-2},0,i_{p-\ell-1},\ldots,i_0)$$

$$+ (-1)^{k_{\ell-1}} j^{k_{\ell-2}} A_{\ell-1}(k_0,\ldots,k_{\ell-2},1,i_{p-\ell-1},\ldots,i_0) \qquad (4.30)$$

$$\cdot e^{-j\frac{2\pi}{N}(k_{\ell-3}2^{\ell-3}+\ldots+k_0)\cdot 2^{p-\ell}}$$

$$k_\nu = 0,1 \quad ; \quad i_\nu = 0,1$$

$$\ell = 1,2,\ldots,p$$

The final transformed values are

$$X(k_{p-1},\ldots,k_0) = A_0(k_0,\ldots,k_{p-1})$$

Note that the indices are reversed. That is, the final data values are not in the correct order in core memory if they are stored by sequence of their index value.

The sine and cosine in the complex exponential values can be generated in advance and stored in a table recursively by the method of Eq. (4.3) where instead of π/m, the quantity $\pi/2^{\ell-1}$ appears in the 2x2 matrix. Keeping in mind that complex arithmetic operations are being performed, one sees that for each element of each A_ℓ, one complex multiply-add operation is required plus some bookkeeping to decide if $(-1)^{k_{\ell-1}} \cdot j^{k_{\ell-2}}$ is positive or negative. Also, one complex multiply-add operation is necessary for the recursive generation of the complex exponential which gives a total of two complex

multiply-add operations. Thus, since there are $N = 2^p$ elements in each A_ℓ and $\ell = 1, 2, \ldots, p$, the total number of operations is

$$T = 2N\log_2 N = 2pN \quad \text{complex operations} = 8pN \quad \text{real operations}$$

since $\log_2 N = \log_2 2^p = p$. It can be reasonably assumed for time estimates that the basic necessary multiply-add loop requires 12 machine cycles plus an additional cycle to account for the sign of $(-1)^{k_{\ell-1}} \cdot j^{k_{\ell-2}}$. For the example used previously, $N = 10,000$.

Assume for here that $N = 2^{14} = 16,384 > 10,000$. Then the total time for the Fourier transform is

$$T = 8 \times 14 \times 10^4 \times 13 \times 1.4 \times 10^6 = 14.56 \text{ sec}$$

An additional 2N operations are required to obtain the raw power spectrum and an assumed $m \cdot 20 = 2 \times 10^4$ of a 41 point smoothing arc and 1000 lag are used. Thus,

$$T' = 4 \times 10^4 \cdot 12 \times 1.4 \times 10^{-6} = .67 \text{ sec}$$

and

$$T + T' = 14.56 + .67 = 15.2 \text{ sec}$$

4.3 DIRECT SPECTRAL COMPUTATIONS

The preceding sections have discussed the efficient computation of correlation functions. Although correlation functions are at times of interest for their own sake, they are often obtained merely as an intermediate step in computing the power spectral density function. In the past, the Fourier transformation of the autocorrelation function was the most efficient method known for obtaining the power spectral density function. However, recent developments in digital filtering techniques have shown that in certain cases it can be almost as efficient and in other special cases more efficient than the correlation function method. The correlation function is obtained from the spectral density function by an inverse Fourier transformation.

In certain cases, advantage can be taken of the physics of the problem which suggest possible speedups in the computational procedures. Variable bandwidth methods are one possibility which is discussed.

4.3.1 Basic Filtering Procedure

Certain basic relations involved in filtering concepts will be reviewed here. First of all, a linear system or operator L is one which fulfills the requirements

$$L[a\,x_1(t) + b\,x_2(t)] = a\,L[x_1(t)] + b\,L[x_2(t)] \qquad (4.31)$$

where $x_1(t)$ and $x_2(t)$ are considered as input variables. A physical linear system can be characterized by a frequency response function or weighting function. A weighting function is defined as the response of the system to a unit impulse function. This weighting function will be denoted by $h(t)$ and can be shown that the output $y(t)$ of a system is then given by the following equation if the input is $x(t)$.

$$y(t) = \int_{-\infty}^{\infty} h(t - \tau)\, x(\tau)\, d\tau \qquad (4.32)$$

The value of the frequency response function H(f) at the frequency point f is a complex number whose absolute value (gain factor) $|H(f)|$ is the change in amplitude of a sine wave which is operated on the system and whose argument (phase factor) $\phi(f)$ is the change in phase or equivalently, the time delay experienced by the sine wave when passing through the system. It can be shown that the frequency response function is the Fourier transform of the weighting function. That is,

$$H(f) = \int_{-\infty}^{\infty} h(t)\, e^{-j 2\pi f t}\, dt = |H(f)|\, e^{-j\phi(f)} \qquad (4.33)$$

It is the shape of the frequency response function gain factor which is of primary concern in designing a spectral analysis. One ideally wants a square, so-called bandpass, filter which will restrict the frequency output to a narrow frequency band of predetermined choosing. That is, $|H(f)|$ should have roughly the form indicated in Figure 4-2. In the digital case, the weighting function will actually be implemented and applied to the input x(t) in order to obtain the output y(t) which for spectral analysis will now be a sharply bandlimited function. The discrete formula for the output would then be

$$y_j = \sum_{i=-p}^{p} a_i x_{i+j} \qquad (4.34)$$

For a so-called recursive (feed back) filter, the output is obtained as a

function of both previous values of the output and input. Equation (4. 34) then becomes

$$y_j = \sum_{i=0}^{p} a_i x_{j-i} + \sum_{k=1}^{q} b_k y_{j-k} \qquad (4.35)$$

Recursive filters are often more efficient computationally than the moving linear combination such as indicated in Eq. (4. 34). However, moving linear combination filters can be constructed with perfect zero phase shift characteristics whereas recursive filters cannot.

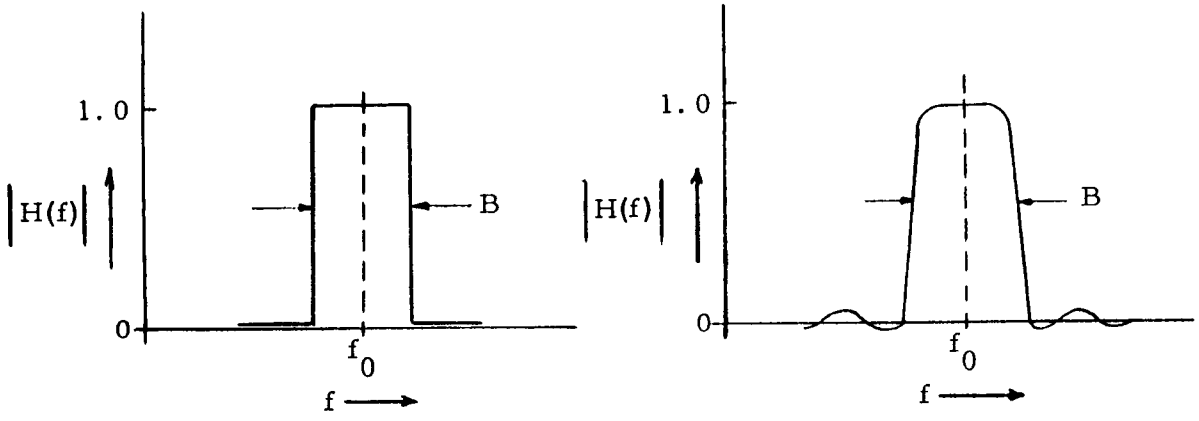


Figure 4-2. Spectral Analysis Filter

Various methods exist for obtaining coefficients a_i in Eq. (4. 34). See, for example, Reference 17 or Reference 18. One method might be of merely specifying the desired frequency response characteristic $|H(f)|$ (gain factor) in a numerical form and then computing the discrete Fourier transform which gives the discrete values of the weigting function h(t). It is these

discrete values of the weighting function which then become the weights a_i in Eq. (4. 34). Other methods exist for evaluating the weights a_i and b_k in Eq. (4. 31) which will not be discussed here but which are available in the cited literature.

The spectral computational procedure employing a direct filtering approach for the power spectral density proceeds as follows:

1. With an appropriate weighting function, compute filtered values of x_i restricted to bandwidth B_a at center frequency f. Let this output be denoted by $X_i(B_a, f)$, $i = 1, 2, \ldots, N$.

2. Compute the variance normalized to unit bandwidth of the sequence $X_i(B_a, f)$, $i = 1, 2, \ldots, N$. Namely,

$$G_{xx}(f) = \frac{1}{NB_a} \sum_{i=1}^{N} X_i^2(B_a, f) \qquad (4.36)$$

3. Select new filter weights with center frequency $f + B_a$ and width B_a. Repeat Steps 1 and 2.

For cross-spectral density functions, a quadrature filter (90^o out of phase) must be applied in order to obtain the imaginary component. Instead of squaring one obtains averaged cross products (the covariance) between the two signals x(t) and y(t). The procedure is

1. With appropriate identical weighting functions compute filtered values of x_i and y_i restricted to a common bandwidth B_a and common center frequency f. Also shift y_i in phase by 90^o. Denote these outputs by $X_i(B_a f)$, $Y_i(B_a, f)$ and $Y_i^*(B_a, f)$, $i = 1, 2, \ldots, N$ where the asterisk indicates the 90^o phase shift.

2. Compute the covariance normalized to unit bandwidth of each of the two pairs of sequences.

$$C_{xy}(f) = \frac{1}{NB_a} \sum_{i=1}^{N} X_i(B_a, f)\, Y_i(B_a, f) \tag{4.37}$$

$$Q_{xy}(f) = \frac{1}{NB_a} \sum_{i=1}^{N} X_i(B_a, f)\, Y_i^*(B_a, f) \tag{4.38}$$

3. Select new filter weights with center frequency shifted by an amount B_a and of bandwidth B_a. Repeat Steps 1 and 2.

This procedure is roughly the one presently employed in biomedical data analysis at UCLA. There it has been found that although for a single record transforming the autocorrelation function is the most efficient procedure, when one wants to compute cross-correlation and cross-spectral density functions among a group of input records numbering more than four, this direct filtering procedure becomes more efficient. This is due to the fact that parallel processing procedures are more easily implemented with the direct filtering approach than with the correlation function approach. It has been found that it is simpler to operate on the data sequentially with direct filtering. With one pass through the data, all spectra and cross-spectra can be more conveniently and more efficiently obtained than they could by attempting to compute the correlation functions.

The term parallel processing refers to approximately simultaneous computations. For example, one can compute an autocorrelation function as data is read into the digital computer to eliminate any record length restrictions. To illustrate this, suppose an autocorrelation function with

m = 5 lag values is to be computed. All lag values can be determined by reading only 2(m+1) = 12 data points at a time into the computer storage.

The process proceeds roughly as follows:

1. Read in data points $x_1, \ldots, x_6$

2. Compute and accumulate products:

$R_{xx}(0)$	$R_{xx}(1)$	$R_{xx}(2)$	$R_{xx}(3)$	$R_{xx}(4)$	$R_{xx}(5)$
x_1^2	$x_1 x_2$	$x_1 x_3$	$x_1 x_4$	$x_1 x_5$	$x_1 x_6$
x_2^2	$x_2 x_3$	$x_2 x_4$	$x_2 x_5$	$x_2 x_6$	
x_3^2	$x_3 x_4$	$x_3 x_5$	$x_3 x_6$		
x_4^2	$x_4 x_5$	$x_4 x_6$			
x_5^2	$x_5 x_6$				
x_6^2					

The above products are then summed by column to obtain partial sums to eventually make up points on the correlation function.

3. Read in data points $x_7, \ldots, x_{12}$, save $x_1, \ldots, x_6$

4. Compute and accumulate the products:

$R_{xx}(0)$	$R_{xx}(1)$	$R_{xx}(2)$	$R_{xx}(3)$	$R_{xx}(4)$	$R_{xx}(5)$
					$x_2 x_7$
				$x_3 x_7$	$x_3 x_8$
			$x_4 x_7$	$x_4 x_8$	$x_4 x_9$
		$x_5 x_7$	$x_5 x_8$	$x_5 x_9$	$x_5 x_{10}$
	$x_6 x_7$	$x_6 x_8$	$x_6 x_9$	$x_6 x_{10}$	$x_6 x_{11}$
x_7^2	$x_7 x_8$	$x_7 x_9$	$x_7 x_{10}$	$x_7 x_{11}$	$x_7 x_{12}$
x_8^2	$x_8 x_9$	$x_8 x_{10}$	$x_8 x_{11}$	$x_8 x_{12}$	
x_9^2	$x_9 x_{10}$	$x_9 x_{11}$	$x_9 x_{12}$		
x_{10}^2	$x_{10} x_{11}$	$x_{10} x_{12}$			
x_{11}^2	$x_{11} x_{12}$				
x_{12}^2					

These products are added to the previous partial sums by column.

5. The next set of six data values would be read into the storage space occupied by the first six. By proper programming one computes the appropriate cross products and accumulates the partial sums as the process goes along so that when the data has been read, all points on the correlation function will have been determined. One can fairly easily generalize this to handle an arbitrary number of variables if the data is multiplexed. That

is, denote the variables by $x_1(t)$, $x_2(t), \ldots, x_p(t)$. The N digitized values of $x_1(t)$ are denoted by $x_{11}, x_{12}, x_{13}, \ldots, x_{1N}$. For the remainder of the variables, the data is then arranged on the digital input tape in the sequence:

$$x_{11}, x_{21}, \ldots, x_{p1} \; ; \; x_{12}, x_{22}, \ldots, x_{p2} \; ; \; x_{13}, x_{23}, \ldots, x_{p3} \; ;$$

$$\ldots \; ; \; x_{1N}, x_{2N}, \ldots, x_{pN} .$$

All auto and cross-correlation functions can then be computed in parallel with the required storage being $3p(m+1)$ cells.

Similar computational principles are employed when direct filtering is used to compute all power and cross-spectra at all frequencies in parallel.

There is one important computational procedure that must be employed before efficiency is gained with the direct filtering method for parallel computations. The idea of decimating the data must be employed. This term refers to the fact that it is not always necessary to compute as many output points of the filter as there are input points. For example, although one has 1000 data input points, it might only be necessary to obtain 200 output points.* See Figure 4-3.

This procedure is justified by the following considerations. The sampling of a random process is usually performed at a fairly high rate in order to avoid the phenomena of aliasing. Although not discussed in technical detail here, this amounts to the fact that it is impossible to distinguish sine waves of multiples of a given sampling frequency. This phenomenon can result in high power at high frequencies being "folded", so to speak,

*Strictly speaking, decimation would refer to dividing into tenths, that is, taking every tenth point. For convenience, it is used here, as in Reference 19, in a general manner, not necessarily restricted to tenths.

across the folding frequency $f_c = 1/2(\Delta t)$ and appearing added in to lower frequency power. Hence, once the data has been sampled at a high enough rate to avoid aliasing, fewer points need be output from the digital filtering operation. One can show that a fewer number of points can be used without any appreciable loss of statistical accuracy. This is due to the fact that the correlated data points contain redundant information. For example, it is claimed on page 45 of Reference 19 that using every fourth data point results in no appreciable loss in accuracy.

For any given application, the amount of decimation which is employed might be adjusted as a function of the aliasing frequency involved and the degree of statistical stability which is desired.

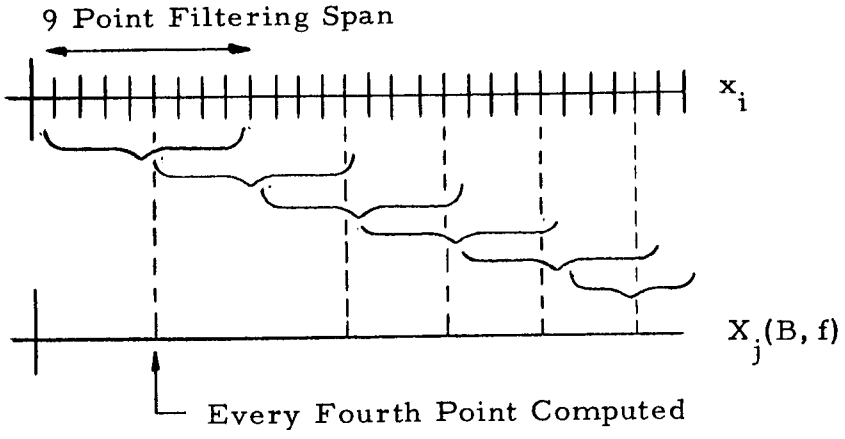


Figure 4-3. Illustration of Decimation

Typical computational times might be as follows. For the previously used example of N = 10,000 and m = 1000, a 41 point moving arc, phase-less filter might be necessary to obtain suitably sharp cutoff characteristics. This requires 20 multiply-add operations per output point. Assuming a 4 to 1

decimation, 2500 points must be squared and averaged to obtain each power spectral density point. If m/2 power spectral density estimates are computed at the spacing $B_a = 1/m\Delta t$, then the total number of multiply-add operations is:

$$\text{Number of Operations} = \left(\frac{N}{4} \cdot 20 + \frac{N}{4}\right) \cdot \frac{m}{2}$$

Assuming that the basic 12 machine cycle multiply-add sequence used for the correlation computation times would apply here, the time estimate for the example is

$$T = 2500 \cdot 21 \cdot 500 \cdot 12 \cdot 1.4 \times 10^{-6} \text{ sec} = 441 \text{ sec}$$

Recursive filters with sharp cutoffs consisting of 7 weights are reasonable. They have the disadvangtage of not having zero phase shift but are computationally efficient. The time estimate for a 7 point filter is

$$T = 2500 \cdot 8 \cdot 500 \cdot 12 \cdot 1.4 \times 10^{-6} \text{ sec} = 168 \text{ sec}$$

Thus, the recursive filter method can be about as efficient as the correlation transformation method.

4.3.2 Use of the Sin x/x Filter

The definition of a frequency response function of a particular form results in a weighting function of an extremely simple form which can simplify the digital filtering operation. One can then apply repeated combinations of this expecially simple weighting function to obtain other more complicated frequency response function forms. A frequency response function given by the equation

$$H(f) = 4\pi fT \frac{\sin 2\pi fT}{2\pi fT} \tag{4.39}$$

has for a weighting function

$$h(\tau) = 1 \quad , \quad -T < \tau < T \tag{4.40}$$

Hence, the discrete weights a_i obtained for this weighting function would all be unity and the filtering operation applied to the input data would not require multiplies and adds, but merely adds. This, of course, results in a more efficient computing procedure.

The penalty paid for this computational efficiency is the fact that the frequency response function shape for Eq. (4.39) does not approach very closely to the ideal indicated in Figure 4-2. The shape of the frequency response function H(f) of Eq. (4.39) is given roughly in Figure 4-4.

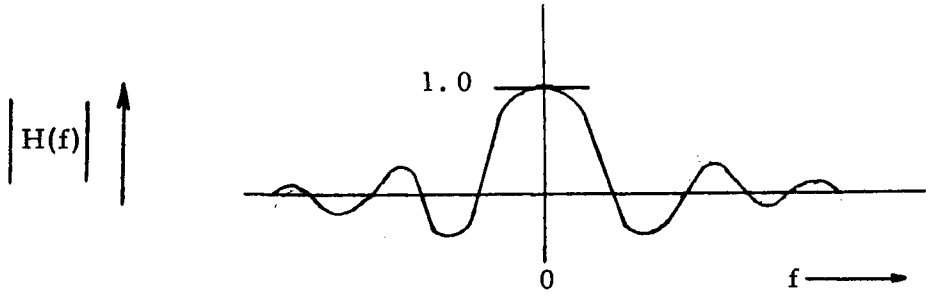


Figure 4-4. Filter Frequency Response of Sin x/x Form

Although Eq. (4.39) in its original form does not give frequency response functions of satisfactory shape, when one takes powers one obtains more useful filters. That is, one defines frequency response functions by the equation

$$H_t(f) = 4\pi\, fK \left(\frac{\sin 2\pi fK}{2\pi fK}\right)^t \tag{4.41}$$

One then employs the well-known theorem which says that multiplication in the frequency domain is equivalent to convolution in the time domain and vice versa. Thus, the multiplication of the frequency response functions to obtain higher powers results in repeated application of the unit weighting function. The weighting function $h_t(t)$ would be given by the formula

$$h_t(t) = \int_{-\infty}^{\infty} h_{t-1}(t)\, h(t - \tau)\, d\tau \tag{4.42}$$

In practice, the modified convolution operation is obtained by reapplication of the unit weighting function to previously filtered data.

There is an additional necessary operation which reduces the computational efficiency of this procedure. In order to maintain this specific simple weighting function (frequency response function), the filter must be kept at zero center frequency. This means that it is required to "demodulate" or "heterodyne" the data in order to map it into a neighborhood of zero frequency. This can be accomplished by the transformation

$$x'_t = e^{j\omega_0 t}\, x_t \tag{4.43}$$

This operation must, of course, be eventually undone by application of an appropriate factor to the final spectral density estimates. This approach has been employed in practice at Princeton University although they have found in general that the autocorrelation transformation procedure is faster. The demodulate-filter method is sometimes more convenient in that the type

of filter used in the spectral analysis can be easily varied by merely taking a higher power of H(f). The direct filter procedure is usually more convenient in obtaining certain higher order spectral density functions which are involved in complicated nonlinear problems.

In this discussion, as in others concerning direct filtering computational methods, the procedures are directly applied to either cross spectra or power spectra. One follows the over-all procedures indicated earlier in Section 4. 3. 1.

This method can have efficient application to problems where only a specific section of the frequency band is of interest. That is, only a small portion of the power spectrum is of interest. The heterodyning of the data down to the low frequency band allows reduced sampling rates to be taken advantage of to reduce the amount of computation. Direct timing comparisons are not meaningful here due to the strong dependence on the freedom to choose specific frequency bands.

4. 3. 3 Single Tuned Filter Methods

It was mentioned in Section 4. 3. 1 that recursive filters are often implemented with more computational efficiency than corresponding moving average filters. A filter which does not have an extremely desirable frequency response function shape, but does have a very simple recursive form, is the so-called single tuned filter or the frequency response function corresponding to a single degree-of-freedom linear system. This frequency response function has a shape roughly as indicated in Figure 4-5.

If x(t) is an input to a single degree-of-freedom linear system, then the output y(t) is given by the extremely simple recursive form

$$y_i = x_i + a_1 y_{i+1} + a_2 y_{i-2} \quad , \qquad i = 1, 2, \ldots, N \tag{4.44}$$

$$a_1 = 2 \exp(-\omega_0 \zeta \Delta t) \cos(\omega_0 \Delta t \sqrt{1 - \zeta^2})$$

$$a_2 = -\exp(-2\omega_0 \zeta \Delta t)$$

$\omega_0 = 2\pi f_0$ = natural frequency

ζ = damping ratio (controls sharpness of response peak)

Δt = sampling interval

This filter might be employed in a variable bandwidth procedure such as is discussed in the following section. The natural frequency f_0 of the filter would be varied to give one a sequence of spectral density points as indicated in Figure 4-6.

In practice this procedure would be implemented by filtering data with a filter located at frequency point f_m, then filtering at frequency point f_{m-1} and subtracting the two outputs. This output is proportional to the power within a fairly reasonable frequency band corresponding to the difference of these two frequency response functions. Although this procedure is promising from the computational efficiency point of view, there are several problems that must be resolved before it is practical to employ. First of all, the shape of the frequency response function is not at all ideal. One must evaluate this non-ideal shape in terms of bias in the resulting spectral density estimates and also in terms of the variability of the spectral density estimates.

The question of bias error has not been thoroughly resolved and no digital application of this procedure is known to the author. Even though the procedure might not be suitable as a highly refined spectral density analysis, it might be very useful as a rough quick-look type of analysis. The single tuned filter approach can be implemented with a two weight recursive filter.

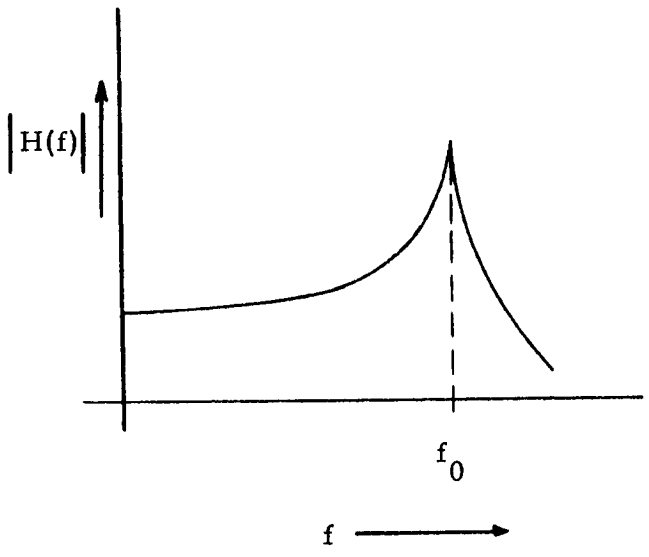


Figure 4-5. Frequency Response Function of Single Tuned Filter

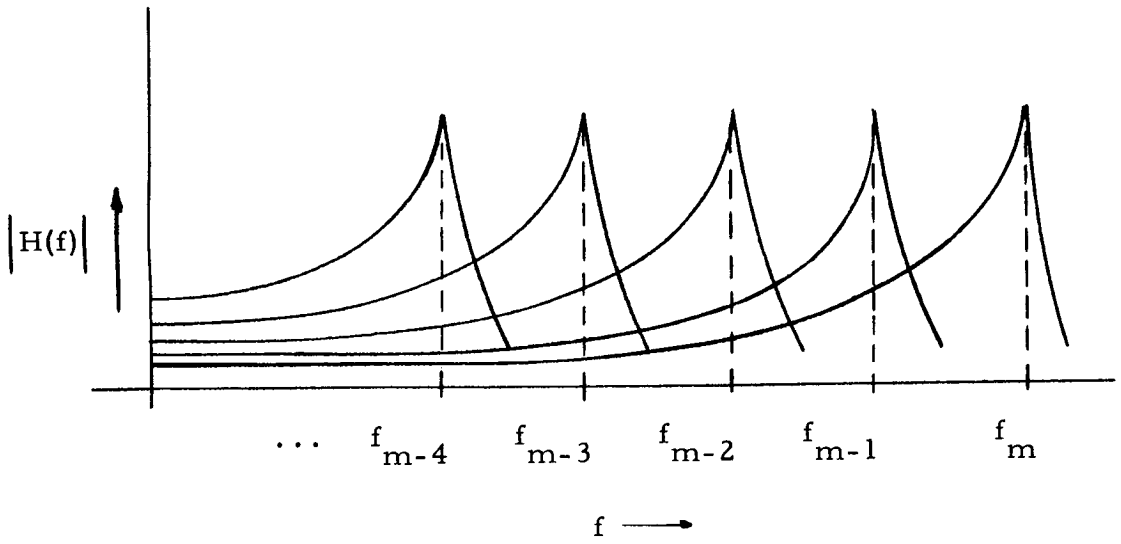


Figure 4-6. Sequence of Single Tuned Filters for Spectral Analysis

The additional add operation needed would require two more machine cycles for each pair of 12 machine cycle multiply-add loops, or a total of 13 cycles per loop. Hence, the time estimate for the N = 10,000 , m = 1000 example is

$$T = 2500 \cdot 3 \cdot 500 \cdot 13 \cdot 1.4 \times 10^{-6} \text{ sec} = 136.5 \text{ sec}$$

4.3.4 Variable Bandwidth Procedures

In all of the discussion up until this time, except for Section 4.3.3, it has been tacitly assumed that the spectral density function has been computed over a frequency band subdivided into a contiguous set of frequency intervals of equal widths. That is, a 20 Hz resolution bandwidth subdivision of the range 0 to 1000 cycles would be 0 to 20 Hz, 20 to 40 Hz, 40 to 60 Hz, etc. This is a proper frequency subdivision for many applications. However, there are certain applications, for example, those of vibration analysis of structures, where a variable bandwidth analysis is more meaningful. This is based on the rationale that the bandwidth of a structural response tends to increase with frequency. Hence, if one wants to resolve two structural response peaks at high frequency,one can get along with a wider frequency analysis than at low frequencies. Thus, a logarithmic type of subdivision of the frequency axis is often suitable. Realizing this fact can result in certain efficiencies in the computing procedure merely by reducing the number of points which are computed on the spectral density function. These topics are discussed in some detail in Reference 20 and will be briefly reviewed here.

Although the exact subdivision of a frequency band into variable widths is in itself a somewhat complicated procedure, one can proceed in the following manner. Suppose that the interval 0 - 1000 Hz is to be divided into R intervals according to an arithmetic progression law with smallest frequency

bandwidth Δ. That is, the width of the ith band is defined as

$$\Delta f_i = i\Delta \tag{4.45}$$

Then Δ may be obtained as (page 5, Reference 18)

$$\Delta = \frac{2(f_{max} - f_{min})}{(R+1)R} = 4.8 \tag{4.46}$$

The results for a 20 interval subdivision of 0 - 1000 Hz are shown in Table 4.2.

Table 4.2. Twenty Interval Variable Bandwidth Subdivision

i	Δf_i	interval endpoint	i	Δf_i	interval endpoint
1	4.8	0	11	52.4	261.9
2	9.5	4.8	12	57.1	314.3
3	14.3	14.3	13	61.9	371.4
4	19.0	28.6	14	66.7	433.3
5	23.8	47.6	15	71.4	500.0
6	28.6	71.4	16	76.2	571.4
7	33.3	100.0	17	81.0	647.6
8	38.1	133.3	18	85.7	728.6
9	42.3	171.4	19	90.5	814.3
10	47.6	214.3	20	95.2	904.8
					1000.0

Note in this case a resolution varying between 4. 8 Hz and 95. 2 Hz is obtained from 20 frequency points. If, for example, the bandwidth of 20 Hz was used throughout, then 50 points would have been required. If a bandwidth of 5 Hz has been used which roughly corresponds to the minimum bandwidth obtained here, then 200 points would be required.

If a direct filtering procedure is used in obtaining the spectral density estimates, then the amount of computing is directly proportional to the number of spectral points. If this variable bandwidth frequency interval subdivision is a reasonable procedure, significant time savings can be accomplished. Programs presently in use have demonstrated times savings of a factor of 3 or greater.

An alternative approach is to break down the frequency interval into a set of, for example, three sets of points, each set having a different bandwidth. A choice of bandwidths that might be employed over the interval 0 to 2000 Hz is depicted in Figure 4-7.

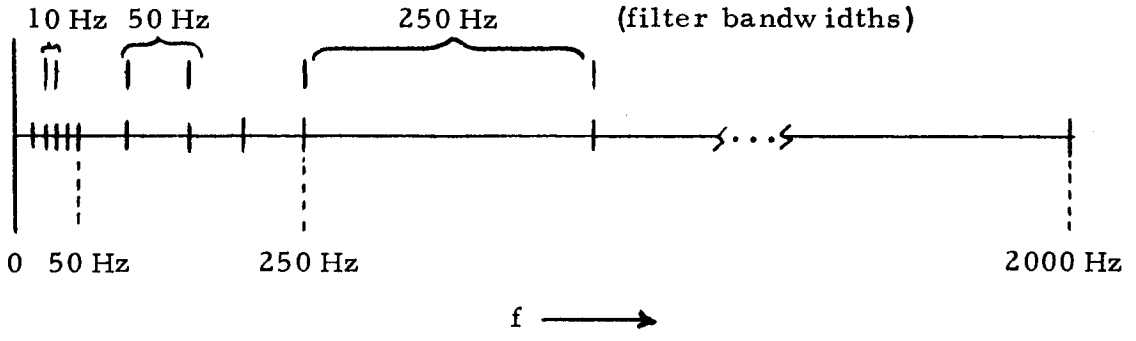


Figure 4-7. Illustration of Variable Bandwidth Procedure

This type of frequency breakdown might be suitable for a crude quick look analysis again. In this case, the spectral density function could be obtained either by a direct filtering method or by a method which involves computing three different autocorrelation functions.

It can be shown that the bandwidth of a spectral analysis is,under certain conditions,roughly equal to

$$B_a = 1/\tau_{max} = 1/m\Delta t \qquad (4.47)$$

Bandwidth is inversely proportional to τ_{max} under any circumstances. Hence, the bandwidth of the spectral analysis can be controlled by the maximum time delay value to which the correlation function is computed. For the three bandwidth frequency subdivision, three different τ_{max} would be chosen,each of which would correspond to the appropriate bandwidth selected.

First the correlation function would be computed out to the smallest τ_{max} value which would correspond to the widest frequency band chosen. This correlation function would be Fourier transformed and the spectral density points computed at the appropriate frequency points desired. Next, in order to avoid aliasing, numerical filtering must be performed which can be followed by decimating. This is logical since there is no point in computing any more correlation points than have been computed on the first correlation function. The same number of points will be computed for the second and third functions, but at wider spacing. This will maintain statistical accuracy roughly equal throughout the whole spectral density function.

In the particular case being illustrated, the decimation would consist of taking every fifth point on the output of the numerical filter since each succeeding bandwidth is one-fifth of the preceding. Likewise, each succeeding maximum time delay is five times the previous. This procedure is repeated two times to give the correlation functions and the maximum bandwidth power spectral density function. It has been estimated in Reference 21 that a quick look spectral density analysis can be performed with a very small amount of computation time when this method is employed. No actual computational experience has been gained with this procedure as of yet, however.

An additional way of subdividing the frequency axis is by octaves. That is, each new interval is double the previous. This has the advantage of reducing to an exactly equal interval subdivision when plotted on a logarithmic scale. A particularly suitable subdivision is that of one-third octave intervals. This has a great advantage in acoustical and vibration applications. This is

due to the fact that certain analog type analyzers have been in existence for many years which employ this subdivision and hence display of results in this format is familiar to many people. However, there is not necessarily any physical justification to the use of the 1/3 octave bandwidth.

As an example, a recursive filter has been developed at Douglas Aircraft Company with approximately linear phase shift with filter gain factors as indicated in Figure 4-8.

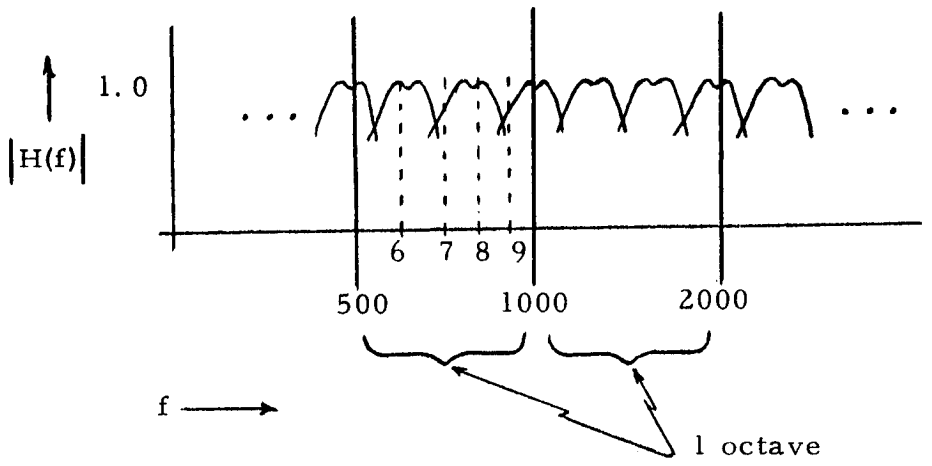


Figure 4-8. Filter Frequency Responses for 1/3 Octave Spectral Analysis

Note each octave is covered by three filters. Centering filters at 500 Hz, 1000 Hz, 2000 Hz, etc., is arbitrary and not absolutely necessary. This type of spectral analysis can be performed relatively quickly due to the minimum number of points necessary to cover a broad frequency range. However, its bandwidth becomes quite broad quickly and it might find its largest application in the "quick look" type analysis.

Experience has recently been gained at Aerospace Corporation with the use of a variable bandwidth power spectral density program. Although it is difficult to compare directly with a constant bandwidth procedure, time reductions of a factor of 4 have been experienced employing bandwidths which are acceptable to the data analysis engineer. Thus, for the example, the time would be

$$T = \frac{160}{4} = 40 \text{ sec}$$

4.4 A SPECIAL DIGITAL STATISTICAL CORRELATION SYSTEM

This system is basically a general purpose high speed, medium scale, digital computer with a special attachment to allow very high speed correlation (and other similar) computations. The machine has a basic machine cycle of 1.75 μsec which makes it slightly slower than an IBM 7094. However, a special "multiply-add box" which is attached to the basic computer is capable of performing the multiply-add loop in this same time of 1.75 μsec.

The multiply-add box has its own separate core memory of several thousand (exact amount unknown) 16-bit words. The multiplication is performed on 16-bit words; however, the accumulator is 32 bits for avoiding overflows. Time estimates have been made of an 800 point correlation function based on a sample size of 8000 data points as requiring less than 13 seconds. This should be about 20.3 seconds for the $m = 1000$, $N = 10,000$ example. This is extremely fast, almost ten times the speed of a typical large scale computer for basic correlation and spectrum computations. The ability to handle large amounts of data is facilitated by a disk file storage device to provide relatively high speed access to large amounts of data points.

Note that the computational procedure followed by the multiply-add box system is a straightforward evaluation of Eq. (4.1). The high speed obtained is strictly by special purpose hardware, not by special computational procedures. However, the high speeds for the specialized correlation multiply-add loop are obtained at reasonable costs as compared to a large scale digital computer. These systems provide some general purpose computing capability in addition to the specialized high speed correlation and spectral computational capability.

4.5 CONCLUSIONS AND RECOMMENDATIONS

A number of methods for the digital computation of correlation and spectral density functions have been described. The optimum choice of a method depends heavily upon the specific application but certain over-all comparisons can be made. Table 4.3 is a summary of estimated computational times for most of the methods discussed in the previous sections.

Table 4.3 Estimated Computation Times for Correlation and Spectral Density Computations

Correlation		Direct Spectral Density	
N = 10,000 and m = 1000			
Basic Method	160.0 sec.	Ordinary Fourier Transform	1600.0 sec.
Factoring	12.9 sec.	Cooley-Tukey Fourier Transform	15.2 sec.
Eight Level Quantization	57.8 sec.	Recursive Filter (7 point)	168.0 sec.
One Bit Quantization	22.2 sec.	Single Tuned Filter	136.5 sec.
Half Clipping	26.7 sec.	Variable Bandwidth	40.0 sec.
Sum of Squares	93.3 sec.	Basic Filtering	441.0 sec.
Special Computer	20.3 sec.		

The time estimates quoted for the correlation or spectral density functions can be considered as the total time required for obtaining both functions. This is because of the fact that the usual numerical Fourier transformation required to obtain a spectrum from a correlation function, or vice-versa, consumes a negligible amount of computational time.

One must exercise care in the interpretation of the times quoted in Table 4.3. In some cases, the analysis bandwidth and record lengths are not exactly the same. For example, the variable bandwidth spectral density computations increase analysis bandwidth with frequency. Thus, one must decide for a specific application as to what would be the necessary analysis bandwidth if a constant bandwidth analysis is used and what would be suitable for a variable bandwidth analysis. Then a more accurate comparison can be made.

The direct Fourier transformation procedures essentially give the narrowest possible resolution. If a correlation approach was employed to obtain the narrowest possible bandwidth, computation time would increase proportionately. For the Cooley-Tukey Fourier transform, the number of data points must be a power of two. Other numbers can be handled but the computation time remains essentially constant until a given power of two is exceeded, at which point the execution time will roughly double. In the case of direct filtering techniques or the ordinary Fourier transform method, the computational time can be reduced by omitting computations for specified frequency points.

Other factors should be taken into account. Many operating programs are available for the standard correlation computations and Fourier transform relations for the spectra. In general, this procedure is straightforward to program. The methods based on extreme quantization sacrifice some information in the data and also can be expected to require additional programming effort. The technique which factors common terms out of the data possibly requires the largest amount of programming effort. The basic filtering methods are generally straightforward to program after a filter characteristic has been decided upon.

Figures 4-9, 4-10, and 4-11 present plots of computational time for various numbers of data points (record lengths) for three values of analysis

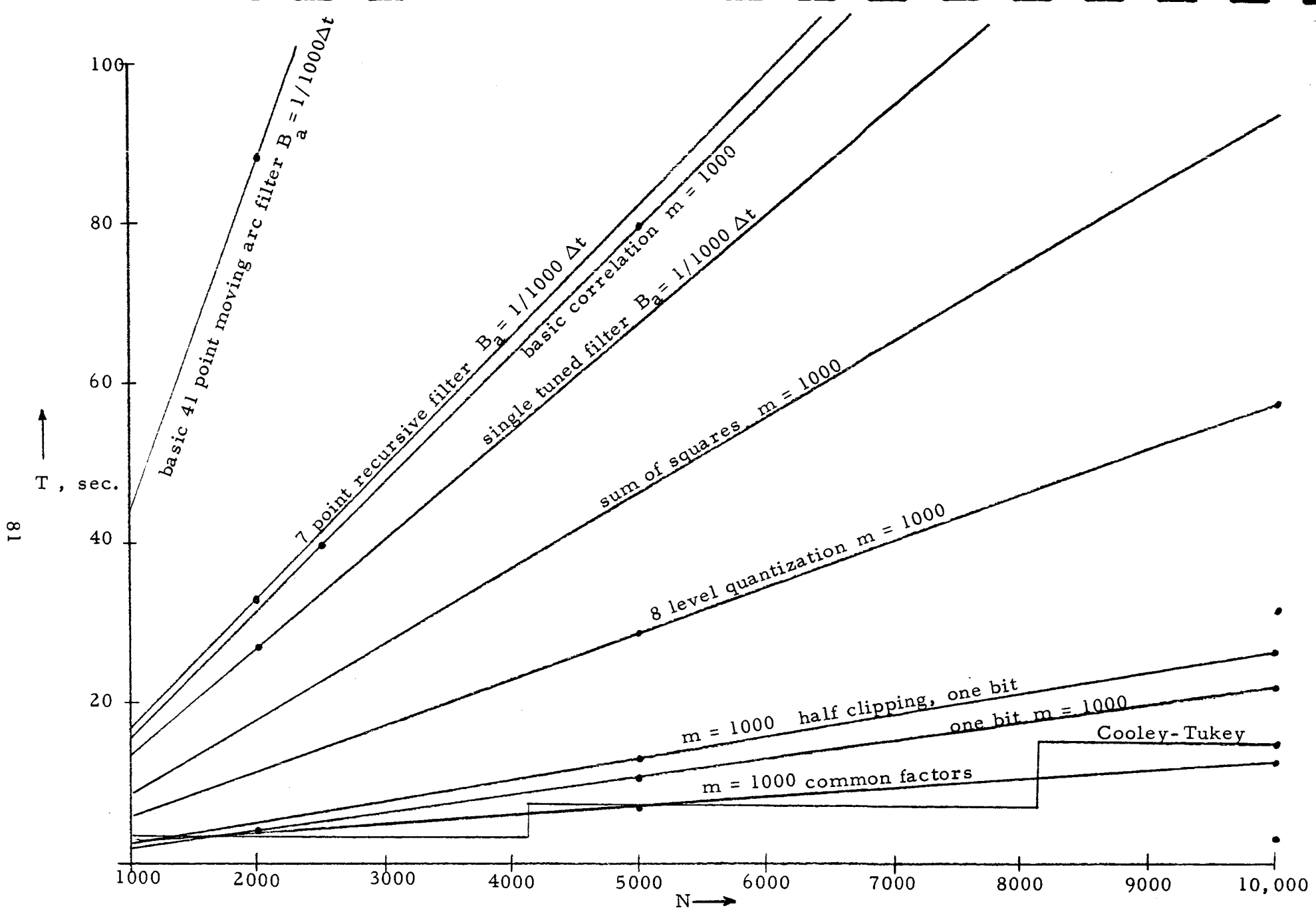


Figure 4-9. Computational Time Estimates as Functions of Number of Data Points
m = 1000 , B_a = 1/1000 Δt

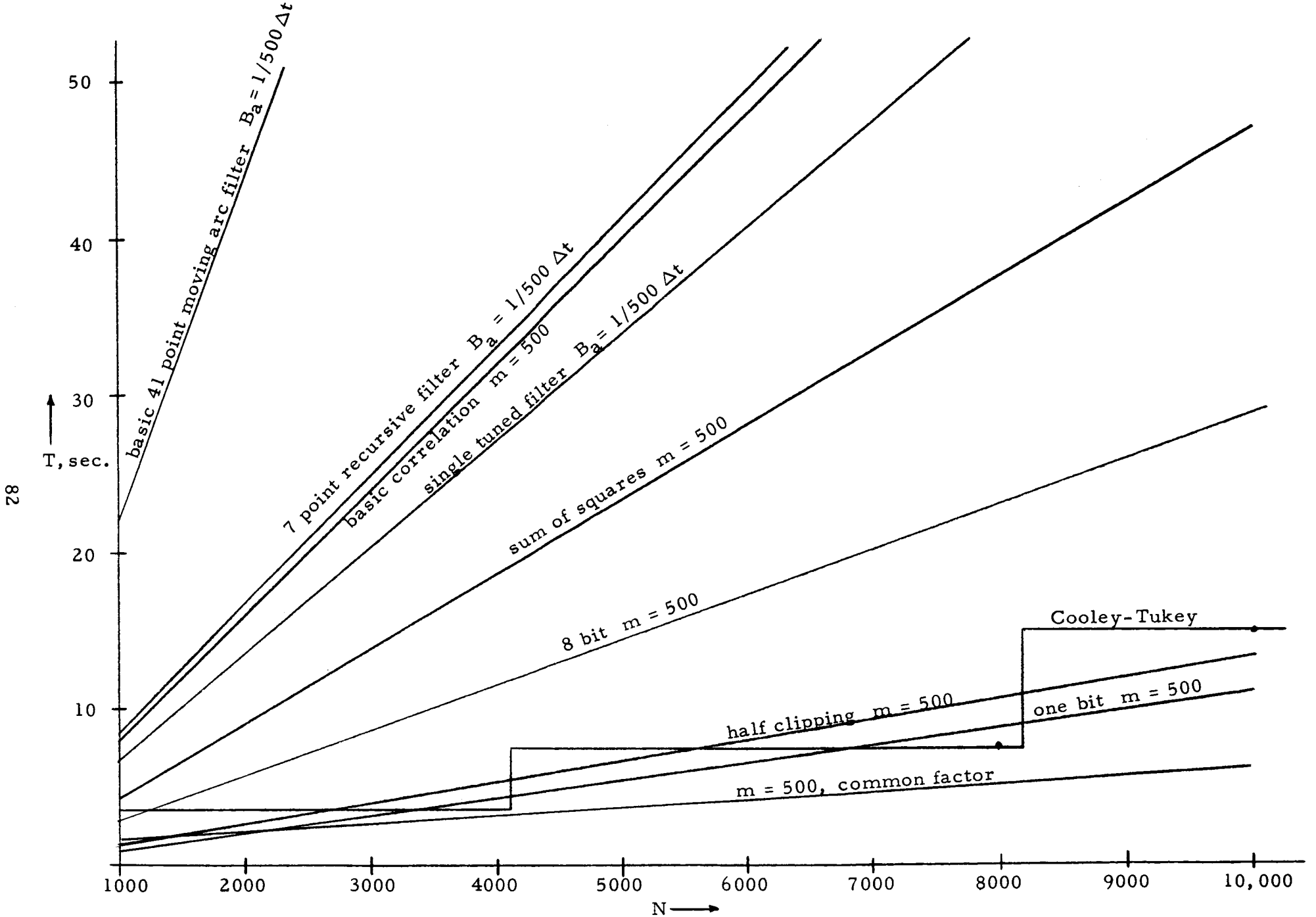


Figure 4-10. Computational Time Estimates as Functions of Number of Data Points, m = 500 , B_a = 1/500 Δt

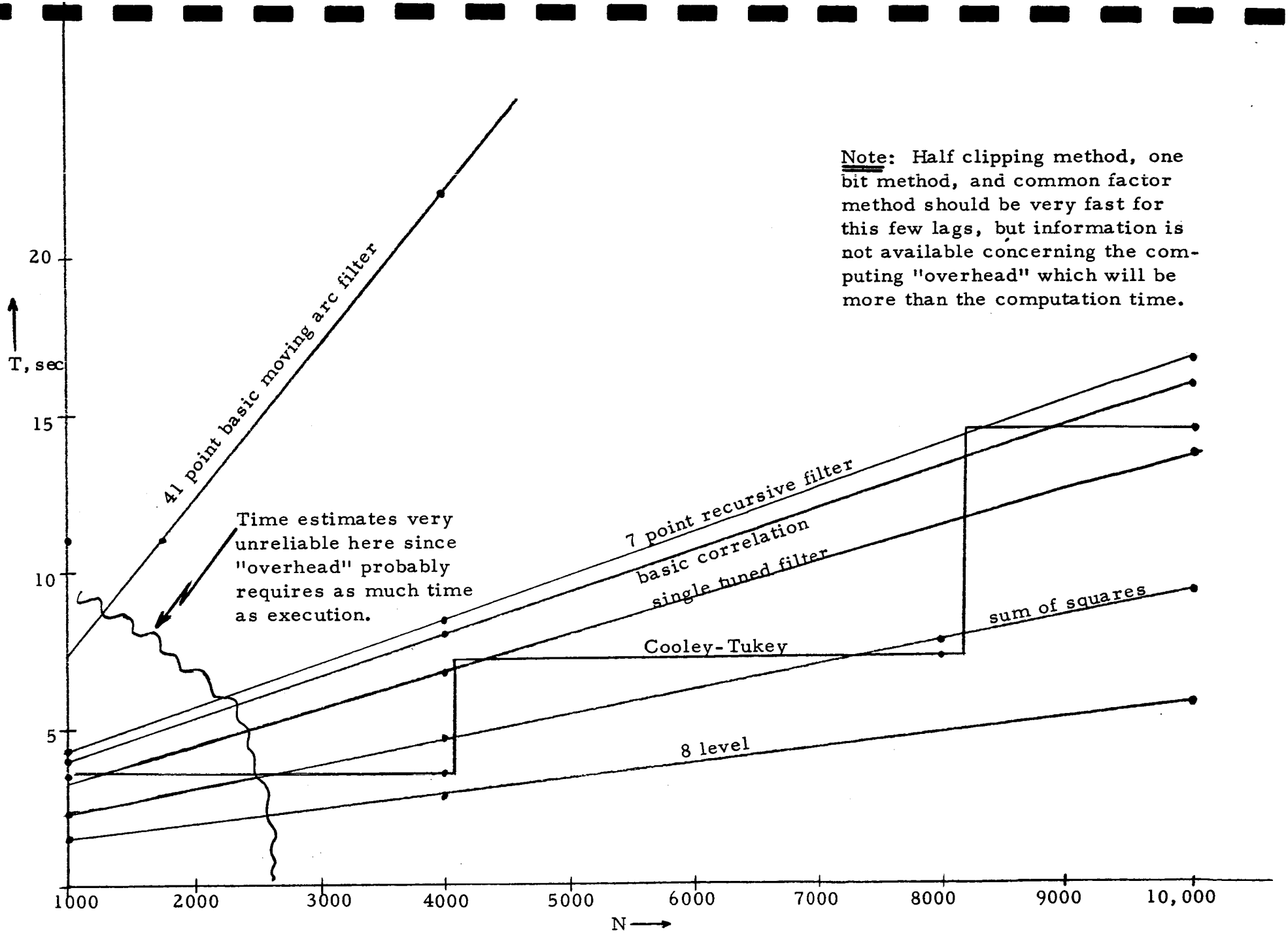


Figure 4-11. Computational Time Estimates as Functions of Number of Data Points
$m = 100$, $B_a = 1/100\ \Delta t$

bandwidth (maximum lag value). Times for the variable bandwidth methods are not included since they can not be directly compared. It would be necessary to plot these times as functions of other parameters. One should keep in mind that these plots are for time estimates and should be used as guidelines only when comparing methods or evaluating computational times. Note that other protions of the program execution time such as data input, and preparation of results for plotting are not included. For relatively small amounts of data these parts of the program require as much, and in many cases more, computational time as the primary computations.

Also, other involved aspects of the program are not included in this evaluation. For example, if very large amounts of data are being analyzed, considerable time might be consumed in transferring data back and forth between storage media such as disc files, magnetic tapes, and basic core memory. For a given specific application, if customized methods are not employed which take advantage of the special requirements of the problem, inefficient processing procedures can easily result.

When all factors are considered, the Cooley-Tukey method of obtaining a Fourier transform is recommended as the best over-all method for a general purpose digital computer. The power or cross-spectral density function can then be obtained by computing variances and covariances of the transforms of the original time histories and then smoothing to obtain desired analysis filter characteristics. The method can be extended to arbitrary (within reason) numbers of data points with modifications to the basic technique. The correlation function would be obtained as a Fourier transform of the spectral density function. An adjustment would be necessary to account for the lag window effect induced by the smoothing of the raw spectral densities.

The programming design necessary for a general program is fairly extensive but the potential time savings is considerable. By far the most

savings are obtained when the programming is accomplished at the machine language level due to the heavy dependence on a binary indexing scheme. If it is necessary to prepare the program in a FORTRAN compiler language then some efficiency is sacrificed but over-all savings still occur. With a direct Fourier transform method, considerable flexibility is available in the selection of smoothing procedures or equivalently, analysis filter shapes. Also note that for this procedure the computations are exact in that there is no depdendence on the form of the data or the quantization. Thus, although extra effort is required for the initial program design and preparation, one obtains an efficient, flexible computational routine for spectrum and correlation computations.

5. HYBRID DATA REDUCTION TECHNIQUES

In this section four specific hybrid computer techniques that have been implemented for the computation of cross-correlation or cross-spectral density functions are discussed. Hybrid computational techniques are gaining popularity because they permit each separate functional operation in a system to be "optimized" by selecting that type of circuitry — analog, digital, or hybrid — that is "optimum" for that particular operation, essentially without regard for the other operations.

5.1 GENERAL PURPOSE HYBRID COMPUTER

One general purpose hybrid computer has been specifically designed for the processing of random data (Reference 22). This device consists of a general purpose digital computer,of fairly high speed but overall moderate capacity,plus special hybrid multiply/divide units that accept one analog input, one digital input, and provide an analog output which is either the product or quotient of the inputs. These hybrid units require 10 μsec to settle to provide 0.01% accuracy in the output. If less accuracy is acceptable in the output, which may well be the case for many applications, then higher rates are possible.

In addition to the basic multiplier unit,a high speed (200 KHz max rate) analog to digital (A/D) converter, digital to analog (D/A) converter, comparator, and multiplexer usually are part of the basic hybrid computing unit. This equipment facilitates the handling of data in either analog or digital form. For correlation and spectral computations, direct evaluation of the correlation function followed by a Fourier transform to obtain the spectrum would most likely be used.

Information on existing programs for this computer indicates that the multiply-add loop should be able to be executed in about 25 μsec, a 40 KHz rate. This speed is possible since the control of the computational loop is performed by the digital computer in parallel with the multiplier. With typical hardware, which has two multiplier units, two points of the correlation function can be obtained simultaneously to double the rate. Thus, one obtains an effective rate of 12.5 μsec which may be compared with an estimated 15.4 μsec for the IBM 7094 operating in the fixed point arithmetic mode.

The computational loop proceeds roughly as follows, assuming the data is in a digital form in core memory:

1. Send x_i to multiplier directly, x_{i+r} to multiplier via D/A converter
2. Obtain $x_i x_{i+r}$ and send to A/D converter and back to main memory to perform accumulation. (Alternatively, some analog averaging device could possibly be attached to accept the products.)
3. While the multiplication is proceeding, the digital computer is performing the necessary address incrementing to provide the next items of data to the multiplier.

The ultimate processing speed,if accuracy is sacrificed,would be limited by the digital computer cycle time of 2.0 μsec. Thus, correlation type processing can be performed at quite high rates with this type of device. This device is also quite flexible in that analog instruments can be inserted anywhere in the analysis chain if desired,or the digital computer can be used to initially convert the raw data into engineering units through the calibrations factors.

5.2 SPECIAL PURPOSE HYBRID COMPUTERS

5.2.1 Time Compression Techniques

There are several commercially available correlation and spectral analyzers whose designs feature the use of time compression techniques.

(For a typical system see Ref. 23.) The spectral analyzers are basically analog instruments to which a digital time compression operation has been added. The advantage of this time compression technique, called DELTIC where delay lines are used for storage, is to provide a speedup of the data and reduction in the analysis time requirements. The result is similar to recording the data on magnetic tape at one speed and then playing it back at a higher speed. The reduction in analysis time is gained as follows:

The scan rate for frequency analysis (Reference 24) is

$$R_s \leq \frac{B}{T_s} \tag{5.1}$$

where

R_s = the scan rate in Hz per second

B = the bandwidth of the analyzer filter in Hz

T_s = the sample record length in seconds

The total analysis time is

$$T = \frac{D}{R_s} = \frac{DT_s}{B} \tag{5.2}$$

where D = the frequency range of the analysis in Hz. If the data is speeded up by a factor of N, T_s is reduced by a factor of N and D is increased by a factor of N.

$$T_s' = T_s/N \quad , \quad D' = ND$$

To maintain the same spectral resolution and uncertainty, the analyzer filter bandwidth can be increased by a factor of N.

$$B' = NB$$

Thus, the scan rate increases by the square of the speedup factor.

$$R'_s = \frac{B'}{T'_s} = \frac{NB}{T_s/N} = N^2 R_s$$

However, the total analysis time is reduced only by the speedup factor.

$$T' = \frac{D'}{R'_s} = \frac{ND}{N^2 R_s} = \frac{1}{N} T$$

The time compression circuit operates by feeding the analog signal to be analyzed to an A/D converter. The output of the A/D converter is fed to digital storage. For analysis, the data is read out of storage in a fraction of the time that was taken to read it in. This data is then run through a D/A converter to obtain the time compressed analog signal for further processing. Two distinct versions of hybrid computers using time compression techniques are available. One employs magnetic core storage and the other uses recirculating delay line storage. One disadvantage of present versions of this type of computer is that they have such limited storage that

frequency analysis is performed with BT products of only one. Thus, the uncertainty error (see Section 7) is horrendous unless special techniques are implemented. The technique most frequently used to reduce the uncertainty is to analyze a large number of sample records and average the results.

Two commercially available correlation analyzers employ this time compression technique. One has magnetic core storage and the other has delay line storage. The correlator using magnetic core storage works similarly to the frequency analyzers. The analog data is digitized, time compressed, converted to analog form, and analyzed in analog circuits. An analog multiplier and an analog integrator are used. The time delay operation is performed during the readout of the time compressed data while it is still in digital form.

The delay line correlator performs all of the analysis operations on the digitized data. The analog data is digitized and read into storage. Then it is read out of storage at a higher rate than it was read in. Time delaying occurs during the readout process. This data is input to a digital multiplier whose output in turn goes to a digital integrator.

5.2.2 Delta Modulation Techniques

Delta modulation is a simple form of pulse code modulation. The output of the modulator consists of a series of constant magnitude, constant width pulses. Only the polarity of the pulse is variable, and it is controlled by the sign of the time derivative of the input signal. Figure 5-1 illustrates the operation of a delta modulation system. This delta modulation technique and a modification of it to obtain dc response (delta sigma modulation) are used in hybrid analog computers for function generation, storage and time delay generation. See References 25 and 26.

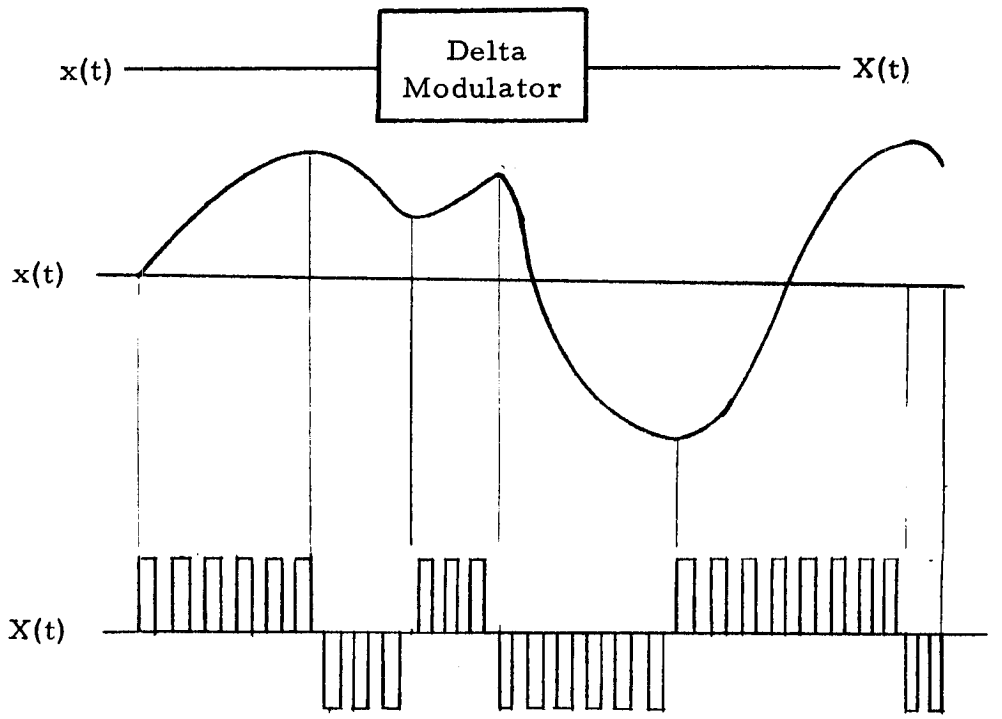


Figure 5-1. Delta Modulation Pulse Code

Function generation is accomplished by creating the basic function $y\,dx$. The pulse output shown in Figure 5-1 is a discrete representation of the time derivative of $x(t)$.

$$\frac{\Delta x}{\Delta t} = c\,u_x \tag{5.3}$$

where

c = a constant of proportionality

u_x = rectangular unity height, constant rate pulses

To reconstruct $x(t)$, one merely integrates the pulse train

$$x(t) = c \int u_x \, dt = \int \frac{\Delta x}{\Delta t} \, dt \tag{5.4}$$

To multiply two data signals x(t) and y(t), one modulates x with $\frac{\Delta y}{\Delta t}$ and sums with y modulated by $\frac{\Delta x}{\Delta t}$. This sum is then integrated.

$$xy = \int \left(x \frac{\Delta y}{\Delta t} + y \frac{\Delta x}{\Delta t} \right) dt \tag{5.5}$$

The block diagram for implementing a correlator based on Eq. (5.5) is shown in Figure 5-2 .

The actual correlator constructed on this principle used several simplifications to reduce the required circuitry (Reference 27). The block diagram of the circuit used is shown in Figure 5-3. The x(t) data signal is converted into the previously described simple pulse code. The pulses are then delayed in a shift register and multiplied against the analog signal y(t). The multiplication is very simple since all that need be accomplished in the multiplier is to switch the polarity of y(t) according to the polarity of the time derivative of $x(t - \tau)$ as represented by the pulses coming out of the shift register. The product is then integrated to obtain the correlation function.

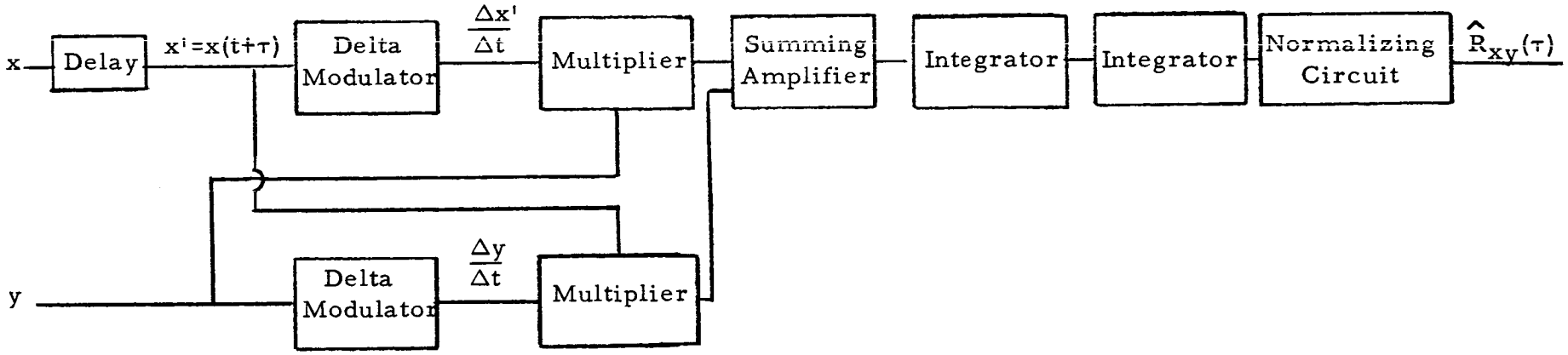


Figure 5-2. Delta Modulation Correlator

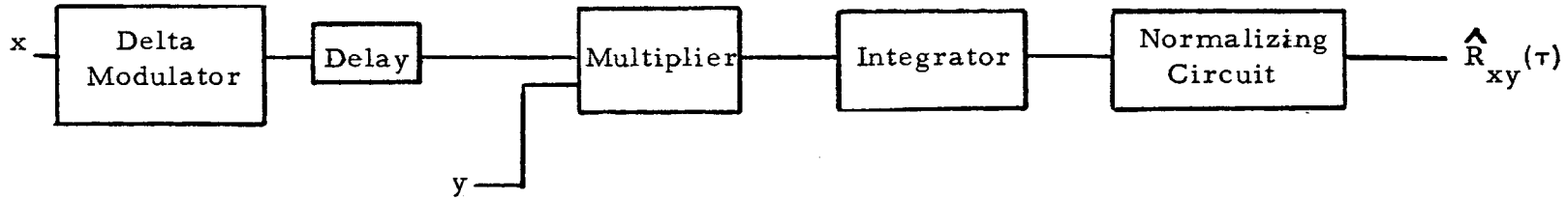


Figure 5-3. Actual Delta Modulation Correlator Block Diagram

5.2.3 Analog Sampling Technique

One commercially available hybrid correlation analyzer (Reference 28) operates by taking analog samples of the two signals to be correlated. The sampling circuits are controlled by a common clock pulse generator. The time delay is created by delaying the clock pulse to one of the two sampling circuits. The sampling circuit outputs feed into a time-division multiplier (see Section 3) and the output of the multiplier is connected to a simple RC integrator. The block diagram of this type of analyzer is shown in Figure 5-4.

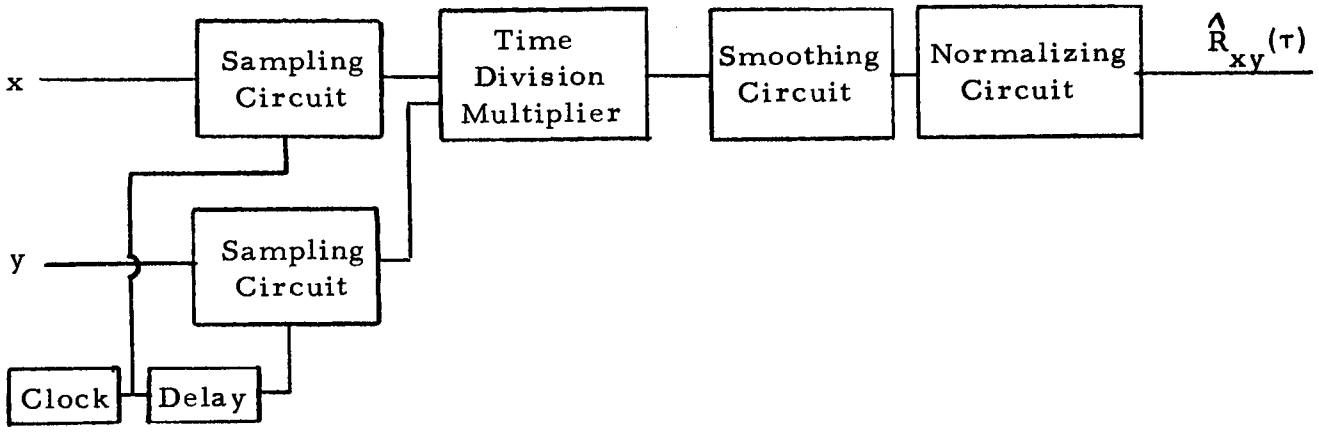


Figure 5-4 Analog Sampling Hybrid Correlator

This computer approximately implements Eq. (2.3), the equation for the computation of a correlation function from discrete data. The major difference, as can be seen below is that τ is continuously variable. (In the actual circuit, the products are also modified by the weighting function of the smoothing circuit.)

$$\hat{R}_{xy}(\tau) = \frac{1}{N} \sum_{i=1}^{N} x(i\Delta t)\, y(i\Delta t + \tau)$$

Because of the design of the analyzer, only analog (pulse height) samples are required. Thus no A/D conversion is necessary and only relatively simple circuits are required to perform all of the operations.

There are several interesting features of this analyzer. First, a clock pulse rate much less than twice the highest frequency is used. From sampling theory, one would expect samples to be taken at a frequency of at least twice the highest frequency of the data to avoid aliasing. This analyzer takes samples at a low rate, throwing away the information between samples much like the decimation scheme discussed in Section 4. The actual limitation on the highest frequency that can be properly analyzed is then one of how closely in time that the two data channels can be sampled (pulse rise time and jitter). This limitation effectively sets the sampling rate in terms of usual sampling techniques.

Because this analyzer uses time division multiplication it is fairly easy to simultaneously compute the cross correlation functions between one channel and several others. Data samples in the common channel drive a pulse width modulator. These width modulated pulses are then used to sample the height of the other data channels. Theoretically there is no limit to the number of channels that can be sampled. The area out of the sampling circuits in each data channel is proportional to the product of the amplitude in that channel (pulse height) and the amplitude of the common channel (pulse width).

The time delay is created by varying the relaxation time in a monostable multivibrator. The clock pulse triggers the monostable and the trailing edge of the monostable output is used to pulse a sampling circuit in the second data channel. Since the clock pulse also triggers a sampling circuit in the first data channel, the second data channel is delayed in time relative to the first by the amount of time it takes the monostable flip-flop to operate. This time is varied by changing a feedback resistor.

5.3 CONCLUSIONS AND RECOMMENDATIONS

The performance of the general purpose hybrid computer discussed in this section compares very favorably with the performance of modern high-speed general purpose digital computers. In fact, from an operational viewpoint, it is probably more convenient than a general purpose digital computer because this hybrid computer can work directly from an analog tape and there is not the need to generate large quantities of digital tapes.

Of the special purpose hybrid computers, the analog sample type of correlator is recommended as the best presently available. This machine has frequency and time delay ranges that are compatible with both the wind tunnel and flight test program requirements. Because it has a single time delay value that is scanned over the τ range, it does require an appreciable time for analysis. This disadvantage is partially regained through this machines ability to simultaneously compute three correlation functions (one data channel common to all three computations). The delta modulation correlator is not being manufactured commercially. The time-compression type of hybrid analyzers are excellent for high resolution spectral analysis and could potentially be the best of the hybrid analyzers if they had much higher A/D conversion rates and greater storage capacity. However, in their present configurations their performance falls short of the requirements of both the flight and wind tunnel programs.

6. APPLICABLE RADAR AND COMMUNICATION SYSTEM TECHNIQUES

6.1 INTRODUCTION

This section presents the results of a survey of both the radar and communications fields to determine special techniques used in these fields that might be useful for satisfying either or both of the dynamic pressure measurement requirements described in Section 1. Only those ideas that appear applicable to the dynamic pressure measurement problem are discussed. This survey was conducted because it was hoped that the techniques used for analyzing the broad basebands available at typical radar and communication frequencies would provide a solution to the problem of analyzing a large number of broadband dynamic pressure channels.

In this section the system requirements are redefined in terms of total storage and total capacity (bits and bit/sec respectively). Three high capacity storage techniques, magneto-optical recording, cathode ray tube, and high speed magnetic disc recording are described. Complete cross-correlation systems employing magnetic drums, film, and electron storage tube devices are discussed. Finally, a comparison is made of the techniques considered.

6.2 SYSTEM REQUIREMENTS

The use of a scale model for wind tunnel dynamic pressure measurements gives rise to data requirements significantly different from those needed for a full scale telemetered flight test. The two sets of data requirements are as follows:

1. Wind Tunnel Test Data
   - (a) Signal bandwidth - 100 KHz
   - (b) Number of channels - 50 (nominal)
   - (c) Test duration - 5 to 30 sec

2. Telemetered Test Data

   (a) Signal bandwidth - 10 KHz

   (b) Number of channels - 250 max (perhaps as few as 12)

   (c) Test duration - 60 to 120 sec

The total storage and capacity implied by the above requirements is conveniently expressed in bits and bits/sec, respectively, (without precluding the possibility of using analog storage). For a signal bandwidth of W Hz, a minimum sampling rate of 2W analog samples per second would be required to preserve all the signal information in a given channel. If the test duration is T_D, the minimum number of samples per channel would be

$$n_s = 2\,T_D W \quad \text{samples/channel}$$

If each of the samples is quantized to b bits, then the required single channel storage would be

$$S_c = 2b\,T_D W \quad \text{bits/channel} \tag{6.1}$$

The corresponding minimum channel capacity (bits/sec) is

$$C_c = \frac{S_c}{T_D} = 2bW \text{ bits/sec} \quad \text{single channel} \tag{6.2}$$

Finally, if there are a total of N_c signal channels, the total storage required for the experiment would be

$$S_{total} = 2bN_c T_D W \quad \text{bits} \tag{6.3}$$

$$C_{total} = \frac{S_{total}}{T_D} = 2bN_cW \text{ bits/sec} \qquad (6.4)$$

Experimental studies at M.I.T. have shown that 3-bit signal quantization (8 levels) yields correlation functions in excellent agreement with results obtained using both 8-bit quantization (256 levels) and analog data (Reference 7). For purposes of this discussion 3-bit quantization will be assumed. Hence from Eqs. (6.1) through (6.4), the storage and capacity (per channel and total) for the two cases of interest may be readily evaluated.

Case 1. Wind Tunnel Test Data Requirements

$$W = 100 \text{ KHz} \quad b = 3 \text{ bits} \quad N_c = 50 \quad T_D = 5 - 30 \text{ sec}$$

$$C_c = 2bW = 600,000 \text{ bits/sec single channel}$$

$$C_{total} = N_cC_c = 30 \times 10^6 \text{ bits/sec}$$

$$S_c = 2bWT_D = 3 \times 10^6 \text{ bits}, \quad T_D = 5 \text{ sec} \qquad (6.5)$$

$$= 18 \times 10^6 \text{ bits}, \quad T_D = 30 \text{ sec}$$

$$S_{total} = N_cS_c = 150 \times 10^6 \text{ bits}, \quad T_D = 5 \text{ sec}$$

$$= 900 \times 10^6 \text{ bits}, \quad T_D = 30 \text{ sec}$$

Case 2. Telemetered Test Data Requirements

$$W = 10 \text{ KHz} \quad b = 3 \text{ bits} \quad N_c = 12 - 250 \quad T_D = 60 - 120 \text{ sec}$$

$$C_c = 2bW = 60,000 \text{ bits/sec single channel}$$

$$C_{total} = N_cC_c = 12(60,000) = .72 \times 10^6 \text{ bit/sec} \quad N_c = 12$$

$$= 250(60,000) = 15 \times 10^6 \text{ bits/sec} \quad N_c = 250$$

$$S_c = 2b\,W\,T_D \cong 3.6 \times 10^6 \text{ bits}, \quad T_D = 60 \text{ sec} \qquad (6.6)$$

$$= 7.2 \times 10^6 \text{ bits}, \quad T_D = 120 \text{ sec}$$

$$S_{total} = N_c S_c = 43.2 \times 10^6 \text{ bits}, \quad T_D = 60 \text{ sec} \qquad N_c = 12$$

$$= 900 \times 10^6 \text{ bits}, \quad T_D = 60 \text{ sec} \qquad N_c = 250$$

$$= 86.4 \times 10^6 \text{ bits}, \quad T_D = 120 \text{ sec} \qquad N_c = 12$$

$$= 1800 \times 10^6 \text{ bits}, \quad T_D = 120 \text{ sec} \qquad N_c = 250$$

Comparison of the two cases shows that the wind tunnel test (case 1) requires both a larger single channel capacity (600, 000 bits/sec) and total capacity (30×10^6 bits/sec) than for the telemetered test (case 2). The single channel storage for the 5 sec wind tunnel test (3×10^6 bits) is slightly less than for the 60 sec telemetered test. However,for the 30 sec wind tunnel test,the single channel storage (18×10^6 bits) is two and one-half times that required for the 120 sec telemetered test. Finally, the total storage required for the wind tunnel test (150×10^6 - 900×10^6 bits) is greater than that required for 12 telemetered channels but less than the storage required for 250 telemetered channels.

For the purposes of this report,the discussion of storage techniques will be confined to the requirements imposed by the wind tunnel test (case 1).

As the discussion later in this section will show, an attempt to achieve the single channel capacity of 600, 000 bits/sec (required in case 1) would require the most advanced laboratory developed high density digital magnetic storage techniques (Reference 27). It also follows that the 50

signal channels could not be handled by a single digital recording channel even if multiplexing circuitry could be developed to handle the 30×10^6 bits/sec total required capacity. Multiple recording channels would be required (perhaps as many as 50, i. e., one per signal channel).

One means of increasing single channel capacity in an attempt to store all 50 signal channels in a single recording channel would be to use cathode ray storage tube techniques (Reference 30). The difficulty with CRT storage techniques is that they are total storage limited to at best 4×10^6 bits per tube using the most advanced tube techniques. Hence, if it were a requirement to store the 900×10^6 bits corresponding to the 30 seconds test duration, it would be necessary to use 225 storage tubes. This is clearly unacceptable from a cost point of view.

A third technique which combines the high information capacity of the CRT digital storage technique, but circumvents its total storage limitation, would be to store the CRT image on film. Since the writing time for a single 4×10^6 bit storage tube might possibly be reduced to one second (Reference 31), a minimum of 8 storage tubes would still be required to attain the 30×10^6 bits/sec required for case 1. Each CRT would be filmed once per second for the duration of the test (i. e., 30 photographs per CRT for $T_D = 30$ sec).

The preceding discussion clearly indicates that unless either the total number of signal channels, or the duration of required test data, or both are reduced, the most advanced storage techniques would still require multiple recording channels. Even with multiple recording channels, the total storage would have to be reduced if tape recording (with even the most sophisticated high density techniques) were used, due to the unrealistically long tape lengths which would be required. This will be made clear in the later discussion of storage techniques.

6.3 REDUCTION OF DATA REQUIREMENTS

The total storage requirements for the wind tunnel test data can be significantly reduced since the minimum record length, consistent with good cross-correlation measurements, is much smaller than the 30 seconds test duration. (This of course assumes that the data are stationary.) The minimum record length is influenced independently by the maximum τ shift in $R_{xy}(\tau)$, the permissible variability error, and the lowest signal frequency of interest in the individual channels. For the wind tunnel test, the values of τ_{max} are expected to be in the range $10 \times 10^{-6} \leq \tau_{max} \leq 3 \times 10^{-3}$ sec. Section 7 shows that record lengths of 10 (or more) times the maximum time shift, τ_{max}, yield tolerable bias errors. This rule of thumb yields as a necessary (but not sufficient) condition a minimum record length of 0.03 sec for the largest τ_{max} of interest. Similarly, Section 7 shows that the normalized statistical uncertainty error can be estimated by $\epsilon = \frac{1}{\sqrt{BT}}$ where B is a measure of the bandwidth of the data and $\tau_{max} \leq 0.1\ T$. Estimating B is a problem for peaked spectra, but the spectra of dynamic pressure data is relatively smooth with a high frequency rolloff. The bandwidth of the data channels is 100KHz; so for a conservative estimate, B will be assumed to be 10KHz. Letting $\epsilon = 0.1$

$$T = \frac{1}{B\epsilon^2} = \frac{1}{(10^4)(10^{-2})} = 0.01 \text{ sec}$$

The influence of the lowest signal frequency, f_0, on minimum record length can be approximated by evaluating the error in the estimate of the temporal variance of a simple sinusoid due to finite record length T. Let

$$x(t) = A \sin \omega_0 t$$

where

$$\omega_0 = 2\pi f_0 = \frac{2\pi}{T_0} \tag{6.7}$$

T_0 = the period of the lowest frequency component. The temporal variance (for infinite record length) is obviously $A^2/2$. For finite record lengths, the temporal variance is

$$\sigma^2_{x_T} = \overset{\sim}{x^2}_T - (\tilde{x})^2_T = \frac{A^2}{2}\left[1 - \frac{\sin 2\omega_0 T}{2\omega_0 T} - \frac{2(1 - \cos \omega_0 T)^2}{(\omega_0 T)^2}\right] \tag{6.8}$$

$$= \frac{A^2}{2}\left[1 - \epsilon(T)\right]$$

where the symbol $\sim$ indicates time average.

The error in sample variance is

$$\epsilon(T) = \frac{\sin 2\omega_0 T}{2\omega_0 T} + 2\left(\frac{1 - \cos \omega_0 T}{\omega_0 T}\right)^2 \tag{6.9}$$

If we concern ourselves only with the peak values of the two error terms, then the second term becomes negligible when the record length exceeds

a few cycles. The table below shows e(T) for several pertinent values of record length.

Number of cycles (T/T_0)	e(T)
2-1/8	0.039
2-1/2	0.032
3-1/8	0.025
3-1/2	0.016
4-1/8	0.019
5-1/8	0.015
10-1/8	0.008

Since the lowest signal frequency will, in general, only represent a small fraction of the total signal spectrum, it is clear from the above table that a record length of 3 to 4 cycles of f_0 should be adequate. For a minimum frequency of 100 Hz, a minimum record length of $T_R = 0.04$ seconds is required. Since this value exceeds the record length dictated by the bias and variability error, a required record length of $T_R = 0.04$ second will be assumed throughout the remainder of this section.

Referring to the wind tunnel data requirements (case 1), the single channel and total storage requirements will be reduced respectively to

$$S_c = C_c T_R = 600,000\,(0.04) = 24,000 \text{ bits/channel}$$

$$S_{total} = N_c S_c = 50\,(24,000) = 1.2 \times 10^6 \text{ bits}$$

These storage requirements may be easily attained by a variety of storage devices. However, although the reduction in the record length greatly reduces the total storage requirements, the total storage rate will still remain at

$$C_{total} = 30 \times 10^6 \text{ bits/sec} \quad \text{(i.e., 50 channels each having a storage rate of } C_c = 600,000 \text{ bits/sec)}$$

It is this requirement which places the greatest demand on the recording system (i.e., the need for multiple recording channels). In the following subsection, advanded storage techniques potentially capable of achieving comparable storage rates will be briefly discussed.

6.4 HIGH CAPACITY STORAGE TECHNIQUES

Three recently developed techniques for achieving high capacity storage are outlined below. The discussion below has been abstracted from References 27 through 30.

6.4.1 Magneto-Optical Recording

Studies of saturation digital magnetic recording (Reference 27) have shown that the major limitation in linear storage density on tape is due to the resolution characteristics of the recording medium. In general, in order to achieve large storage capacities (bits/sec) requires transporting the medium past the heads at fast speed, and usually implies the use of disc or drum memories without contact recording, since direct contact is undesirable in high speed applications due to head and coating wear.

The major limitation on linear storage density for drum or disc memories (in which the heads do not contact the coating) is due to the separation of the playback head from the coating material when this separation is approximately 0.0002 inch or greater. This limitation results from the fact that conventional playback heads are sensitive to the magnetic flux changes external to the surface of the coating material. This limitation increases in importance as the resolution characteristics of coating materials improves.

MAGOP (magneto-optic playback head) makes use of a beam of light to read the surface magnetization of the coating material without contacting the coating. Use is made of the Kerr magneto-optic effect, i.e., the optical rotation that results when light is reflected from a surface which is magnetized in the plane of the film and parallel to the plane of incidence. A new electric vector (Kerr component) is created at right angles and in addition to the E vector that would result without magnetization. The two vectors combine in the reflected beam to form a new vector and thus a rotated plane of polarization. The direction of rotation depends upon the direction of surface magnetization. The light reflected from the surface consists of an average light level on which the useful magneto-optically produced signal is superimposed.

A block diagram of the system is shown below (Reference 27, Figure 6).

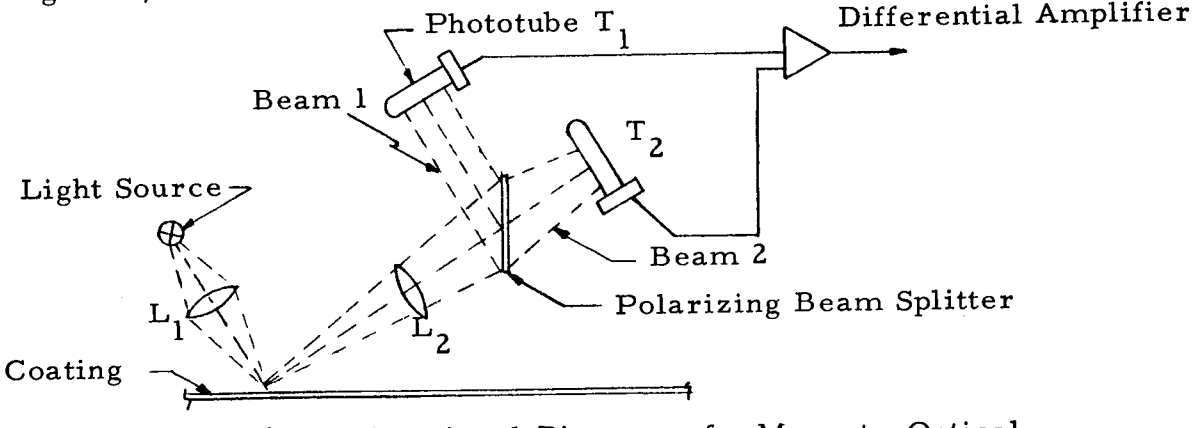


Figure 6-1. Functional Diagram of a Magneto-Optical Playback System

Light from the source is focused by lens L_1 onto the magnetic coating film. A polarizing beam splitter is used to form two beams (polarized at right angles to each other) from the light reflected from the magnetized film surface. Each of the polarized beams may be made to yield a useful magneto-optic signal by proper orientation of the plane of polarization with respect to the plane of incidence. While the signals in the two beams are 180° out-of-phase, the film noise components, produced by variations in the film surface reflection, are in phase. By subtracting the outputs from the two phototubes in a differential amplifier the entire useful magneto-optic signal is retained while film noise (a chief source of deterioration) is canceled. The main contribution to noise is limited then to the photo detector shot noise.

In addition to the influence of the polarization beam splitter on reducing film noise, new film techniques have been developed which allow transverse recording, and the combination of these effects has resulted in the ability to obtain storage densities of 1000-3000 bits/inch using a conventional record head with a 0.5 mil head to coating spacing. On the basis of microscopic observations the recorded magnetization was found to extend only 2 mils beyond the record head so that track densities of 100 track/inch are achievable. Hence surface storage density may be as high as 300,000 bits/inch^2.

Since there is no contact of the coating material, linear recording speeds of 1200 inches/sec are possible, resulting in recording capacities of 1.2 - 3.6 megabits/sec. The technique is applicable to tape, disc or drum recording, although tape would be less desirable due to the high speed indicated.

6.4.2 Cathode Ray Storage Tubes

The cathode ray storage tube principle offers a variety of techniques for the storage and/or conversion of electrical or visual information (see

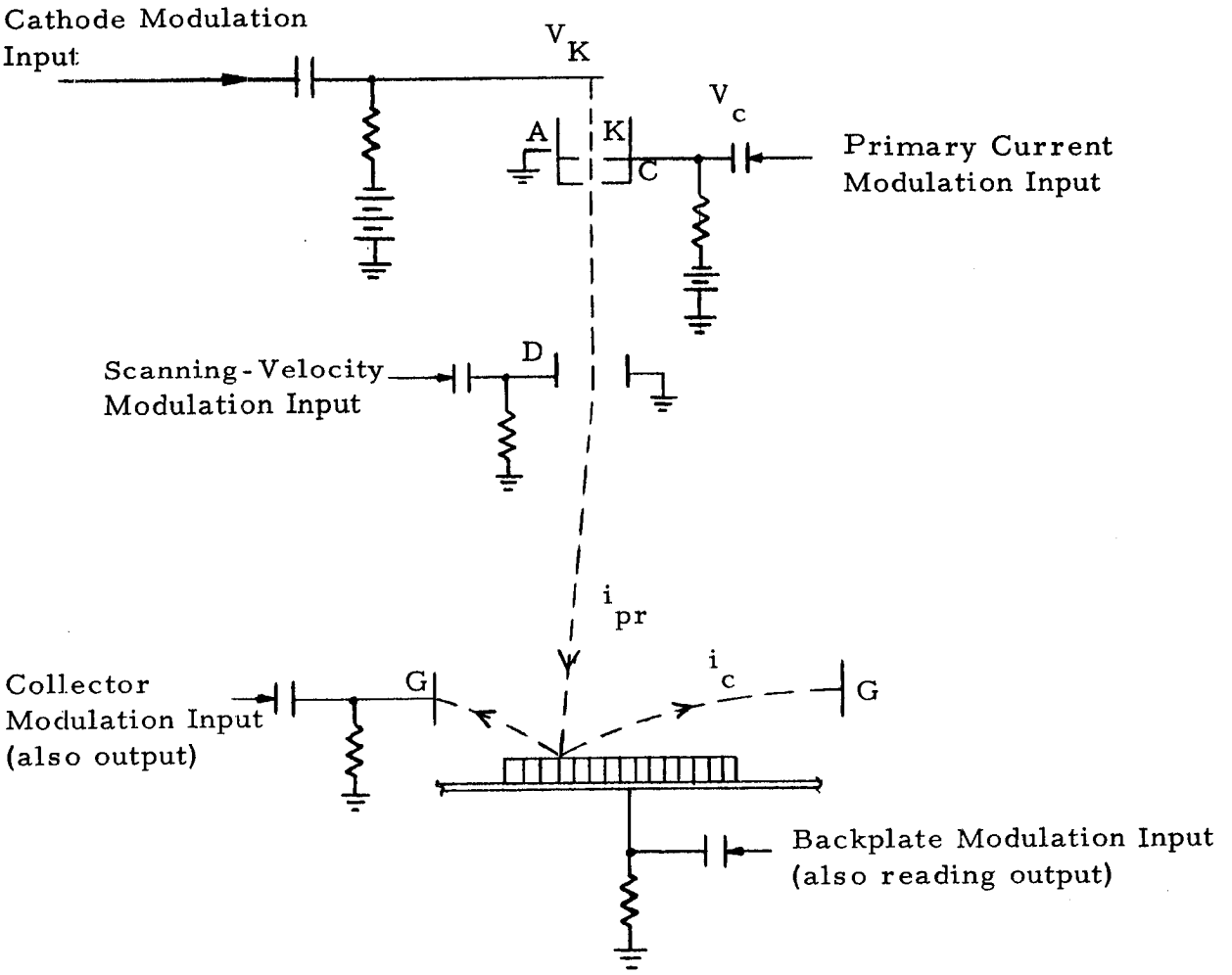
References 30 and 31). In all storage tubes, information is stored in the charge pattern on an insulated storage surface (target) placed between an electron gun and a metal backplate. Each insulated target element comprises a miniature condenser between its front surface and backplate. If the target surface is composed of photo-emissive elements, then writing may be achieved by scanning the target with a beam of light instead of an electron beam.

The writing process consists of establishing a charge pattern on the target elements either by regulating the number of primary incident electrons or by control of the number of secondary electrons leaving the target through the use of a collector electrode. In the reading process the elementary condensers of the target are scanned by an unmodulated primary electron beam giving rise to potential shifts at each element which in turn result in available output signals in either the backplate or collector circuits.

Several of the variety of methods for applying input signals to the storage tube during writing and extracting output signals during reading are indicated in Figure 6-2. (See Figure 9, Reference 30)

The following important operational characteristics are available in various storage tubes (not all in the same tube though):

1. Signals may be stored in analog or digital form.
2. Writing and reading may be performed by the same electron gun.
3. Simultaneous writing and reading can be accomplished by using separate guns mounted at opposite ends of the tube.
4. Readout may be designed to be either destructive or nondestructive.
5. Signal bandwidths of 3 to 6 MHz have been achieved.



Methods of Applying Input Signals to, or Obtaining Output Signals from, a Storage Tube.

K	cathode	T	insulated target
A	accelerating annode	P	backplate
C	control grid	i_{pr}	primary beam current
D	deflection plates	i_c	collector current
G	collector		

Figure 6-2. Functional Diagram of a Cathode Ray Tube

In addition, the following input-output relations can be achieved:

a. Electrical input — Electrical output
b. Electrical input — Visual output
c. Visual input — Electrical output
d. Visual input — Visual output

One of the shortcomings of cathode ray storage tubes has been their limited total storage capability. However, recent advances in the cathode ray art have resulted in tubes which permit the display of 4×10^6 resolvable beam positions (an array of 2000 x 2000) with adequate guard bands to prevent crosstalk (Reference 31). Read and write rates from 4×10^6 bits/sec (assuming digital storage) to perhaps 12×10^6 bits/sec (assuming analog storage with eight distinguishable levels), can be achieved.

6.4.3 High-Speed High-Capacity Magnetic Disc Recording

In general, the achievement of high bit packing densities for noncontact magnetic recording (associated with high-speed high-capacity storage) requires the use of minimum recording medium thickness, minimum transducer to surface spacing, and a well designed pole tip configuration whose magnetic influence upon the recording medium is sharply defined in both the writing and reading process (see Reference 32).

With the advent of air bearing gliding head techniques it has become possible to operate transducers within 200 to 300 microinches from thin recording media.

By combining this technique with improved pole tip design, the Development Laboratory of IBM General Products Division, San Jose, California, has developed disc recording transducers which have achieved linear storage densities of 520 bits/inch with a track speed of 1200 in/sec beneath the

transducer yielding a storage capacity of 625,000 bits/sec. Track densities of 50 track/inch have been achieved.

6.5 CROSS-CORRELATION TECHNIQUES

For purposes of this study it is convenient to initially classify radar and communication system cross-correlation techniques into two broad categories:

a. matched filter techniques
b. nonmatched filter techniques

The reason for this is simple. In radar and/or communication systems it is often the case that one of the two signals to be cross-correlated has a temporal waveform which is known a priori. Starting with this supposition it follows that the cross-correlation of an unknown signal $x(t)$ with a known signal $y(t)$ is obtainable as the output of a linear filter whose impulse response is "matched" to $y(t)$. This result arises from the similarity between the correlation integral and the convolution integral of linear filter theory. The cross-correlation function of $x(t)$ and $y(t)$ is

$$R_{yx}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, y(t - \tau)\, dt$$

Now if $x(t)$ is the input to a linear filter with impulse response $h(\tau)$ then the output $z(\tau)$ is given by

$$z(\tau) = \int_{-\infty}^{\infty} x(t)\, h(\tau - t)\, dt$$

In order to make these two integrals identical all that is required is to set $h(\tau - t) = y(t - \tau)$, i.e., $h(u) = y(-u)$. A filter with this impulse response is said to be matched to $y(t)$. If $y(t)$ begins at $t = 0$ and has a duration T, then the impulse response of its matched filter would not be physically realizable since $h(t)$ cannot be nonzero for negative t. This situation can be corrected by simply adding a delay of T seconds into the filter. The only effect this has is that it delays the cross-correlation function at the output by T seconds. Thus the output of the matched filter contains the entire cross-correlation function for all values of τ.

Though the matched filter technique offers a means of obtaining the entire cross-correlation function as the temporal output of a linear filter, it is not easily made amenable to the problem at hand; namely cross-correlation of two unknown signals. Hence no attempt will be made to describe the various matched filter implementation schemes.

However, there have evolved a number of radar processing techniques which allow for the cross-correlating of a priori unknown signals from two or more independent channels. Several of these techniques, which differ in one or more basic respects, and which are representative of present technology, will now be briefly described.

The major difference in the various cross-correlation systems is in the form of signal storage employed. The three techniques to be considered are

1. magnetic drum storage
2. film storage
3. cathode ray storage tube

Relative signal delays can be achieved by

a. delay lines

b. multiple playback heads on a magnetic drum

c. shifting relative positions of film transparencies

Signal multiplication and integration are normally achieved by using a mixer followed by a narrowband filter. Optical multiplication is achieved by an overlay of two transparencies, on each of which is stored one of the two signals being multiplied. A third technique makes use of an analog electron beam tube multiplier.

6.5.1 Magnetic Drum Storage Pulse Doppler System

The system is based on the pulse-doppler principle where short samples of coherent echo pulses are recorded and stored on a rotating magnetic drum. A block diagram of the entire system is shown in Figure 6-3.

In this technique a replica of a transmitted pulse is stored in a recirculating delay device in which the delay is equal to the transmitted pulse width τ.

A repetitive playback of the stored replica of the transmitted pulse into the mixer M1 and filter F1 gives a complete cross-correlation between each transmitted pulse and the entire signal present at the receiver between one transmitted pulse and the next. Each playback constitutes a "look" at the receive signal in the next successive range gate. Each new transmitted pulse replaces the previous pulse in the recirculating delay device. In the radar problem the desire is to cross-correlate a number of successive transmitted pulses separately with the received signals from each range gate. This is achieved by use of the sampling filter and the storage drum whose rotation is synchronized to the transmitter pulse repetition frequency. For a moving target in any range cell, the sampled bi-polar video output of filter F1 will exhibit the doppler frequency in the form of echo height modulation on successive pulses. Due to the synchronous drum rotation, successive

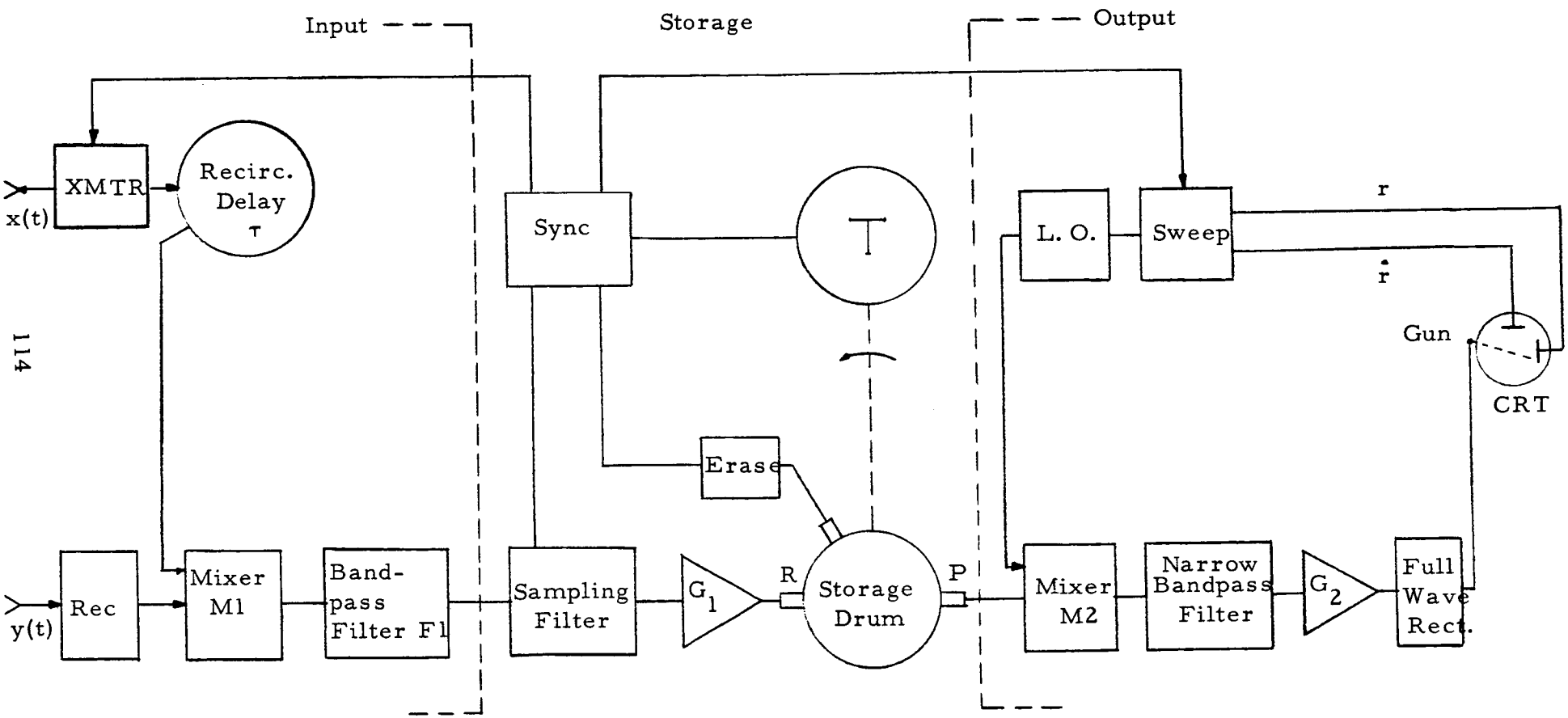


Figure 6-3. Block Diagram of a Pulse Doppler System

pulse sampled returns from a given range gate will be stored adjacent to one another on the drum. The stored pulses consist of a carrier signal on which the doppler frequency appears as amplitude modulation. Since the desired radar information is contained in the modulation, the output from storage has to respond only to the modulation envelope.

If the radar integrates for T seconds with a pulse repetition frequency f_r, the number of samples stored for each range gate are

$$n_s = f_r T$$

Since these pulses are stored in an interval of τ seconds on the drum, the pulse width out of the sampling filter must be

$$\tau_s = \frac{\tau}{n_s} = \frac{\tau}{f_r T}$$

If the pulse width is τ, the number of range gates is

$$m_r = \frac{1}{f_r \tau}$$

The output pulses which are read off in τ seconds will experience a multiplication in doppler frequency of

$$p = n_s m_r = \frac{T}{\tau}$$

To determine the output doppler frequency, the output signal from the drum is mixed in M2 with a local oscillator (L. O.) whose frequency is scanned through the doppler interval required. For each revolution of the drum, all range gates will be "sampled" for one doppler frequency. In succeeding revolutions each range gate is displaced around the drum at the rate of one stored pulse width per revolution and all "sampled" for the next adjacent doppler frequency. In a time T, all range gates will have been displaced by one range gate width and all doppler frequencies up to the pulse repetition rate will have been sampled. At this point the local oscillator will recycle and the process will be repeated.

One final note with respect to the sampling technique. Since the sampling filter samples each range gate at the rate of the p. r. f., the p. r. f. should be chosen at least twice as large as the highest anticipated doppler frequency to ensure getting at least two samples per cycle.

Applicability

If we disregard the output operations employing the swept L. O., the remainder of the system (with modifications) would appear to be applicable to the cross-correlation of arbitrary signals. The major problem areas would be the storage rate (capacity) and the total storage volume limitations of the drum. These questions will be considered in further detail later in this section.

6.5.2 Film Storage Electronic System

A second type of radar cross-correlator which uses film as a storage medium makes use of the known doppler history of the signal returned to a moving radar from a fixed target, as a function of range and azimuth. The return pulses are received coherently and heterodyned down to video with a constant offset frequency. The coherent video is recorded on film using a cathode ray tube as indicated in Figure 6-4.

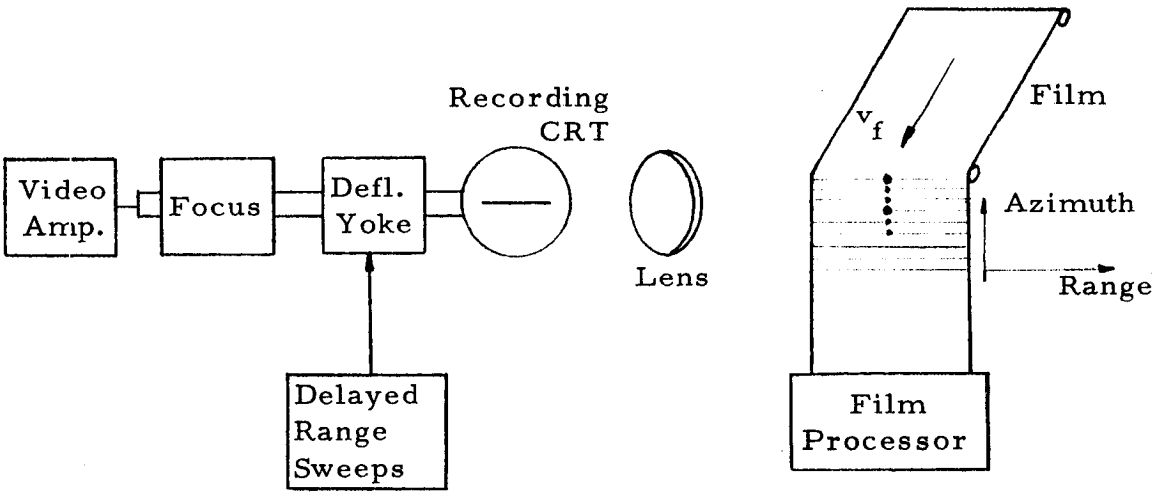


Figure 6-4. Block Diagram of a Film Storage System

For any target the doppler history is contained in the amplitude modulations of the video envelope. This history is recorded as spot intensity on the recording film.

After the film is processed, the raw data is read out orthogonally using a flying spot scanner as indicated in Figure 6-5.

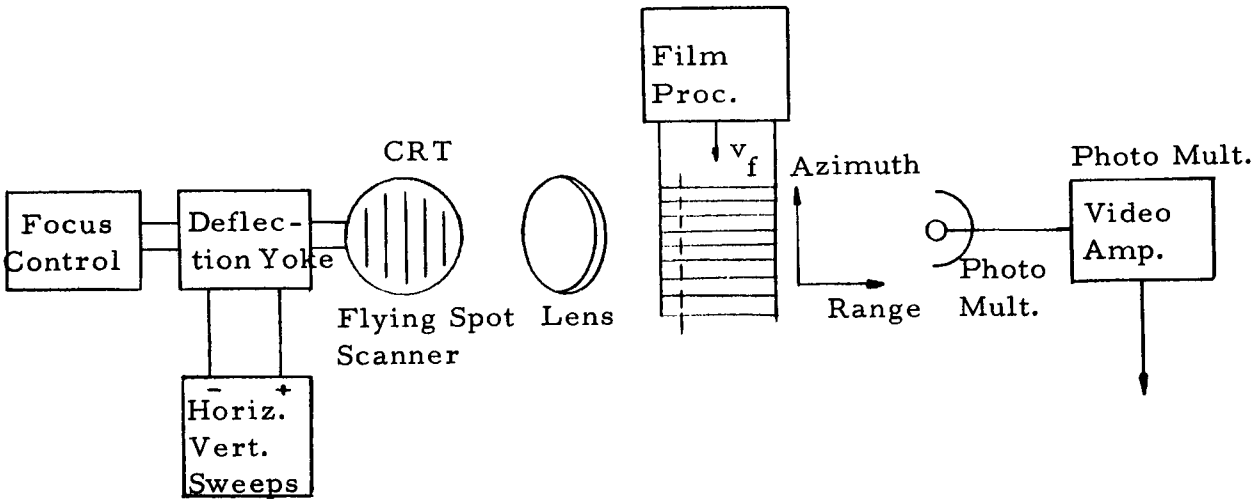


Figure 6. 5. Block Diagram of a Film Readout System

The flying spot scanner raster scans and illuminates the film. The light transmitted through the film transparency contains the signal information in the form of intensity modulation. This intensity modulated light is amplified by the photomultiplier tube.

The output from the photomultiplier-video amplifier is then cross-correlated with a predicted doppler history which is obtained from a swept oscillator as shown in Figure 6-6.

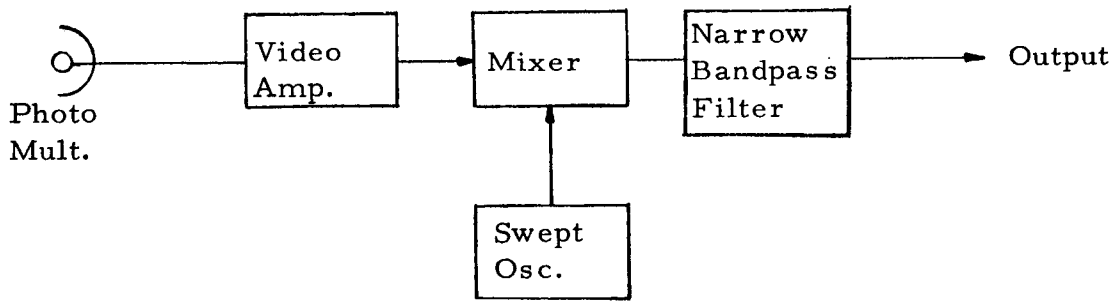


Figure 6-6. Block Diagram of a Correlation System

Applicability

This technique could be applied to the cross-correlation of two unknown signals by replacing the swept oscillator by the delayed output from a second channel which could also have been recorded and read out by the above outlined techniques.

6.5.3 Electron Storage Tube System

A third technique makes use of a cathode ray storage tube to store the products of the two signals to be cross-correlated. Due to the large signal bandwidth (100 KHz) and hence, high data rates involved, it seems reasonable to assume that the signals to be cross-correlated will be stored in analog form. In handling the data from 50 signal channels, multiplexing techniques are called for in an effort to reduce the cost of the storage system. In theory, if each signal has a bandwidth W Hz, then it is only necessary to sample each channel 2W times per second. However, this assumes that the signal will be reconstructed by passing the successive samples (appropriately delayed) through a filter having an impulse response of the form $\frac{\sin(\pi Wt)}{\pi Wt}$. In practive, in order to avoid the implementation problems associated with such an approach, each signal should be sampled at a rate of 4W to 5W samples per second. (References 33 and 23.)

For a sampling rate of 4W and a signal bandwidth of 100 KHz each channel must be sampled at a rate of 400,000 per second. Since present storage tube technology is limited to perhaps 4×10^6 (i.e., 2000 x 2000) resolvable beam positions in a writing time of one second, a single tube should be capable of storing in analog form the multiplexed signals from ten channels. Hence, a total of five storage tubes would be needed to achieve the required storage rate for simultaneously recording 50 signal channels. Each tube has sufficient storage to handle one second of sampled test data concurrently from ten signal channels.

Since as discussed earlier, it is only necessary to store a record length of 0.04 second for each channel, it follows that each storage tube need only utilize 160,000 beam positions (i.e., ten channels, 400,000 samples/sec and T = 0.04) for initial data storage.

A considerable reduction in tube cost can undoubtedly be achieved by reducing the tube storage requirement from 4×10^6 to 160,000 resolvable beam positions. Furthermore, the assumed storage rate of 4×10^6 samples per second would be more easily achieved with fewer resolvable beam positions. While the larger number of resolvable beam positions would increase the tubes' flexibility as a computing device, it would appear that from the point of view of availability, reliability and cost effectiveness the lower tube storage volume is preferable.

Since the signals are stored in sampled analog form, the required cross-correlation (with finite record length)

$$R_{yx}(\tau) = \frac{1}{T - \tau} \int_{\tau}^{T} x(t)\, y(t - \tau)\, dt$$

may be approximated by the summation

$$\hat{R}_{yx}(\tau) = \frac{1}{MW\,(T - \tau)} \sum_{k=\ell=MW\tau}^{MWT} x(t_k)\, y(t_k - \tau)$$

where MW is the number of stored samples per second,

$t_k = t_0 + \frac{k}{MW}$ (t_0 is arbitrarily set equal to zero)

$= t_0 + k\Delta$

and $\tau = \frac{\ell}{MW} = \ell\Delta$

where $\Delta = \frac{1}{MW}$ and ℓ is an integer

For M = 4, W = 100 KHz, and T = 0.04 second there are 16,000 product pairs to be summed for each value of τ. A single storage tube having 160,000 resolvable beam positions could store all the product pairs for ten separate points of the cross-correlation function. A tapped delay line is used in conjunction with a sampler and multiplier to achieve the multiplexed product pairs for each of the ten values of τ. The schematic diagram (Figure 6-7) illustrates the technique for obtaining and storing these product pairs.

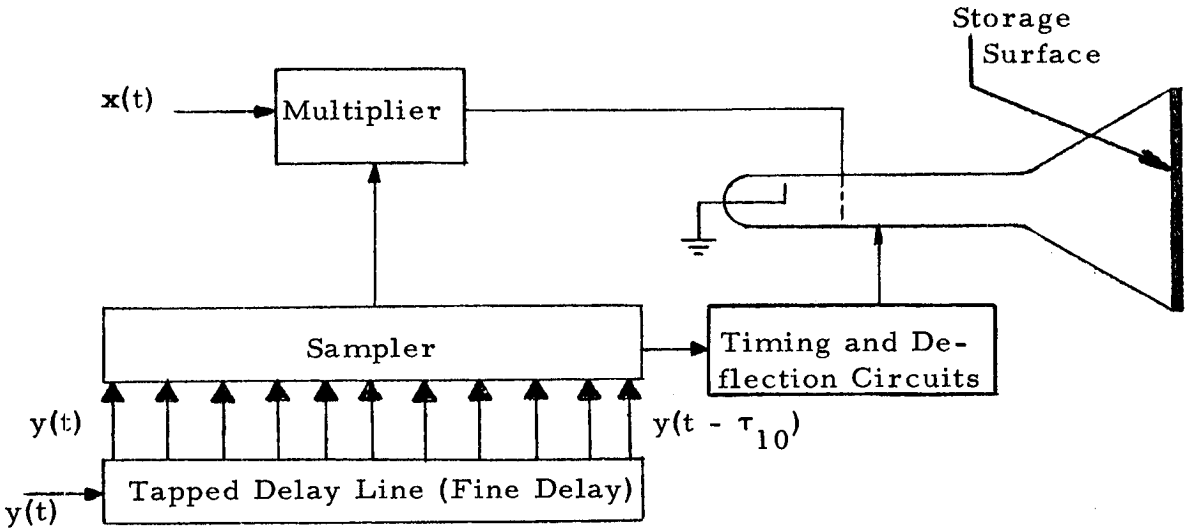


Figure 6-7. Functional Diagram of a Correlator Using an Electron Storage Tube

The signals x(t) and y(t) are assumed to have been previously sampled and recorded. In addition to the fine delay increments provided by the tapped delay line, a gross variable timing delay can be inserted in either channel to change the τ interval on successive replays.

By properly synchronizing the tube beam deflection circuits with the sampling circuits, the 16,000 product pairs for each of the τ delay channels can be stored in 40 columns (400 beam positions per column). In this way the readout and integration of the stored product pairs will be greatly facilitated.

If the stored samples of x(t) and y(t) are played back at the same rates at which they are recorded,then both the sampler and multiplier must be capable of operating at 4 MHz rates. Only analog multipliers can be considered for such operating rates. Several methods of achieving analog multiplication are listed below:

1. Quarter square multipliers (Reference 7 , page 43)
   a. Square law vacuum tube
   b. Beam deflection square law
   c. Diode network
2. Electron beam tube (Reference 34)
3. Time Division (Reference 35)

Methods 1b, 2, and 3 are all capable of achieving the required multiplication rates. Of these, the time division technique (Reference 35) appears to be more compatible and certainly less costly than the electron beam tube multipliers. The time division multiplier, described in detail in Reference 35 , makes use of a Mark/Space modulator, and has achieved accuracies of less than one percent with mean sampling rates of 3 MHz. The use of higher speed transistors would allow the sampling rate to be increased to the required 4 MHz.

Since in this discussion the sampled values of the signals x(t) and y(t) are assumed to be replayed at the same rate as they were recorded, the product pairs for all ten τ values will be stored in a time interval of the

order of 0.04 second. However, there is no reason that the recorded sample values of x(t) and y(t) cannot be played back at a reduced rate. The required rate of multiplication and sampling would be accordingly reduced, thereby resulting in simplified multiplier circuit design and sampler synchronization problems. If the recorded sampling interval Δ is increased to $M\Delta$ on playback then the time delay spacings on the tapped delay must also be increased by a factor of M. The time required to record all the product pairs will likewise be increased by a factor of M.

In either case, once the product pairs are stored, the process of reading out and integrating the product signal for each value of τ can be accomplished by a variety of standard approaches (References 2, 7, and 36). If the product signal is read out in the same time scale as the originally sampled signals, then an analog integrator having a time constant of 10 T could evaluate successive values of $R(\tau)$ in approximately 0.4 second so that 250 points would require 100 seconds to evaluate.

The system as described uses only one multiplier and one integrator. The computation time could be significantly reduced, if need be, by the use of parallel integrator circuits.

6.6 CONCLUSIONS AND RECOMMENDATIONS

In principle the three techniques presented could all be made suitable for the cross-correlation or arbitrary random signals such as those present at the transducer outputs in the wind tunnel measurements of dynamic pressure.

In the magnetic drum approach, it is the limited number of storage positions on the drum circumference rather than the storage rate which necessitates the use of separate recording tracks to handle the required 50 signal channels. Recording sampling rate bandwidths of 6 to 12 MHz appear

to be feasible; however, a practical value for the number of resolvable circumferential pulse positions per track appears to be of the order of 60,000 - 90,000 (Reference 37). Since a signal record length of 0.04 second requires 16,000 analog samples, a single recording track would require 32,000 storage positions for 2 channels and 64,000 storage positions for 4 channels. Assuming a storage density of 1000 pulses per inch, a drum diameter of approximately 20 in. would be required to handle 4 channels on a single track. The 4 channels could be stored by using 4 separate write heads located at 90^{o} intervals around the circumference, or by multiplexing the channels and using a single write head. In the multiplexed case the drum speed would have to be four times as fast to maintain the same sampling rate per channel. The use of separate read and write heads for each channel without multiplexing would appear to be preferable since the τ delays would be varied by a relative rotation of the heads.

As noted earlier, the film storage technique could be adapted to the cross-correlation of two arbitrary wideband signals by replacing the swept oscillator by the delayed output from an independent duplicate film storage channel. The attainable recording bandwidth with such a system should be comparable to that for the cathode ray storage tube provided the total storage does not require mechanical film transport during recording. This is predicated on the assumption that the conventional CRT employed in this technique is of comparable quality to its storage counterpart. The advantage of this technique (when compared to the cathode ray storage tube approach) is that it provides permanent storage. The presence of the optical subsystems would probably make this technique more costly to implement and maintain than either the drum or storage tube approach.

The storage tube technique is clearly applicable, as presented, to the cross-correlation of many arbitrary wideband signal channels. Although each storage tube was assumed to store the sampled signals for only 10

channels, the synchronization of the multiple storage tubes should not be difficult to achieve. In this way the system would maintain the flexibility to cross correlate the signals from any two channels. Although each of the three techniques presented will permit synchronization, the storage tube technique would appear to be superior to either the drum or film storage techniques. On the other hand, with conventional tape recorders, it is not presently feasible to achieve sufficient synchronization to allow the accurate cross correlation at high frequencies of signals stored on separate recorders. The seriousness of this limitation will depend upon the relation between the number of channels which can be multiplexed on a single recorder and the required number of cross correlations.

It should be noted that even with rather coarse quantization (3-bit) and very short record lengths (0.04 second), the three above techniques just barely satisfy the operational requirements. While these techniques could be converted to the analysis of dynamic pressure measurements, it does not appear that they offer any major advantages over existing techniques.

7. MEASUREMENT ERRORS

This section describes the errors that occur in the actual measurement of cross-correlation and cross-spectral density functions. Both the statistical uncertainty errors occuring from analyzing data samples of finite duration and finite bandwidth and the errors associated with practical hardware are covered. The hardware errors apply to all items in the measurement system, from the transducer through the analyzer. Ordinary hardware errors such as those related to frequency response, amplitude linearity, etc. are not discussed. For discussion of these types of errors see References 38 and 39.

Emphasis on the hardware errors is placed on phase errors and those other errors where the previous theory required extension. Detailed examinations of the errors associated with the dynamic phase errors of magnetic tape recorders and the finite size errors of transducers are conducted because it appears that these two items may introduce the largest hardware-related errors into the measurements.

7.1 STATISTICAL ERRORS IN CROSS-CORRELATION AND CROSS-SPECTRAL ESTIMATION

The purpose of this subsection is to present the errors in estimating cross-correlation and cross-spectra caused by finite record lengths, finite bandwidths, extraneous noise, and other effects which are inherent in the estimation procedures. It is assumed throughout that the records being operated upon are samples from stationary, ergodic random processes with zero mean value.

7.1.1 Cross-Correlation Estimation

Let x(t) and y(t) be sample records from two stationary, ergodic random processes. If each record has duration T sec, the sample cross-correlation function is defined by

$$\hat{R}_{xy}(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} x(t)\, y(t + \tau)\, dt \quad , \quad 0 \leq \tau \leq T \tag{7.1}$$

For negative values of τ, the cross-correlation function may be found from

$$\hat{R}_{xy}(-\tau) = -\hat{R}_{yx}(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} y(t)\, x(t + \tau)\, dt \tag{7.2}$$

Therefore we may restrict our attention to positive values of τ since similar error formulas will hold for negative τ.

Upon taking the expected value of both sides of Eq. (7.1), it is seen that

$$E\left[\hat{R}_{xy}(\tau)\right] = R_{xy}(\tau) \tag{7.3}$$

so that $\hat{R}_{xy}(\tau)$ is an unbiased estimate of $R_{xy}(\tau)$. The variance in the estimate is given by

$$V[\hat{R}_{xy}(\tau)] = E[\hat{R}^2_{xy}(\tau)] - R^2_{xy}(\tau)$$

$$= \frac{1}{(T-\tau)^2} \int_0^{T-\tau}\int E[x(u)\, y(u+\tau)\, x(v)\, y(v+\tau)]\, du\, dv - R^2_{xy}(\tau) \tag{7.4}$$

Thus it is seen that the variance of $R_{xy}(\tau)$ involves a mixed fourth moment of x(t) and y(t) which, in general, is quite difficult to calculate. To simplify Eq. (7.4), assume that x(t) and y(t) are Gaussian random processes. Under this condition, Eq. (7.4) becomes

$$V[\hat{R}_{xy}(\tau)] = \frac{1}{T-\tau} \int_{-(T-\tau)}^{T-\tau} 1 - \frac{|\eta|}{T-\tau} \left[R_x(\eta) R_y(\eta) + R_{xy}(\eta+\tau)\, R_{yx}(\eta-\tau)\right] d\eta \tag{7.5}$$

where

$$\eta = v - u$$

$$d\eta = dv$$

If the maximum value of $R_{xy}(\tau)$ approaches the square root of the product of the autocorrelation functions of x(t) and y(t), both evaluated at τ equal to zero $\left(\left|\sqrt{R_x(0)\, R_y(0)}\right|\right)$, then it can be shown (Reference 40) that the variance of $\hat{R}_{xy}(\tau_m)$ approaches

$$V[\hat{R}_{xy}(\tau_m)] \approx \frac{R^2_{xy}(\tau_m)}{B(T-\tau_m)} \tag{7.6}$$

where

τ_m = the value of time delay corresponding to the maximum value of the correlation function

B = the bandwidth of the signals being analyzed (assuming that both signals have the same bandwidth)

In many correlation analyzers, division by the variable time interval $(T - \tau)$ is replaced by division by the record length T. This results in an estimate of $R_{xy}(\tau)$ which is given by

$$\hat{R}_{xy(T)}(\tau) = \frac{1}{T} \int_0^{T-\tau} x(t)\, y(t + \tau)\, dt \tag{7.7}$$

This estimate is biased, however, as can be seen by taking expected values in Eq. (7.6).

$$E[\hat{R}_{xy(T)}(\tau)] = \left(1 - \frac{\tau}{T}\right) R_{xy}(\tau) \tag{7.8}$$

Thus, while the correct functional form is estimated, the amplitude scale is in error.

In certain instruments, the quantity τ is scanned continuously rather than in discrete steps. This means that an additional error will be present in the estimate. If the τ variable is changed at a rate of λ, then the estimate becomes

$$\hat{R}_{xy(s)}(\tau_1) = \frac{1}{T} \int_0^{T-\tau_1} x(t)\, y(t + \tau_0 + \lambda t)\, dt \tag{7.9}$$

where $\tau_1 = \dfrac{\tau_0 + \lambda T}{1 + \lambda}$. Assuming λ is small (which is always the case in practical analyzers) and that λT is equal to or less than the reciprocal of the highest data frequency, then Eq. (7.9) may be written

$$\hat{R}_{xy(s)}(\tau_1) \cong \frac{1}{T} \int_0^{T-\tau_1} x(t) \left[y(t + \tau_0) + (\lambda t)\, \dot{y}(t + \tau_0) \right] dt \tag{7.10}$$

Thus,

$$E[\hat{R}_{xy(s)}(\tau_1)] = \frac{T - \tau_1}{T} R_{xy}(\tau_0) + \frac{\lambda (T - \tau_1)^2}{2T} R'_{xy}(\tau_0) \tag{7.11}$$

Scanning τ increases the magnitude of the variance of $\hat{R}_{xy}(\tau)$. An expression is given in Reference 9, page 280, for the variance of $\hat{R}_{xy(s)}(\tau)$; however, its complexity precludes practical use. But for many cases of interest it can be shown that when the scan rate is small enough to meet reasonable bias error values, the increase in variance due to scanning will be negligible.

Another source of error in correlation measurements is the presence of noise added to the recorded data. Thus, instead of x(t) and y(t) being present at the correlation input, the quantities

$$\begin{aligned} x_1(t) &= x(t) + n_1(t) \\ y_1(t) &= y(t) + n_2(t) \end{aligned} \tag{7.12}$$

are present. It may be safely assumed that $n_1(t)$ and $n_2(t)$ have zero mean and are mutually independent and each is independent of x(t) and y(t). Under these conditions the estimate of $R_{xy}(\tau)$ becomes

$$\hat{R}_{xy(n)}(\tau) = \frac{1}{T}\int_0^{T-\tau} x_1(t)\, y_1(t+\tau)\, dt \tag{7.13}$$

Because of the assumption of independence, it is easily shown that

$$E[\hat{R}_{xy(n)}(\tau)] = \frac{T-\tau}{T} R_{xy}(\tau) \tag{7.14}$$

so that no bias error is generated by the extraneous noise. Without going into the detailed calculations, it may be shown that the variance is given by

$$V[\hat{R}_{xy(n)}(\tau)] = \frac{1}{T}\int_{-(T-\tau)}^{T-\tau} \left(1 - \frac{|v|}{T-\tau}\right)[R_x(v)\, R_y(v) + R_{xy}(v+\tau)\, R_{yx}(v-\tau)]\, dv$$

$$+ \frac{2}{T}\int_0^T \left(1 - \frac{v}{T-\tau}\right)[R_{n_1}(v)\, R_{n_2}(v) + R_x(v)\, R_{n_2}(v) + R_y(v)\, R_{n_1}(v)]\, dv \tag{7.15}$$

Thus, the variance is composed of two distinct parts. One part consists of the sampling error which is due only to the finite record length while the second is the contribution of the two extraneous noise sources. Examination of Eq. (7.15) indicates that, unless the extraneous noise power is small compared to the signal power, the variance of the estimate will be greatly increased. It is difficult to present a quantitative estimate of the increase in the error because the correlation functions of the extraneous noise sources will not be known in general.

7.1.2 Cross-Spectral Estimation

In estimating the cross-spectrum, $G_{xy}(f)$, each record is passed through a narrowband filter to produce outputs $x(t, f, B)$ and $y(t, f, B)$ where f is the center frequency of the filter and B its bandwidth. Since $G_{xy}(f)$ is generally a complex function, its real and imaginary parts must be estimated separately. Thus, denoting the cross-spectrum by

$$G_{xy}(f) = C_{xy}(f) - j\,Q_{xy}(f) \tag{7.16}$$

the estimates of $C_{xy}(f)$ and $Q_{xy}(f)$ are given by

$$\hat{C}_{xy}(f) = \frac{1}{BT}\int_0^T x(t, f, B)\; y(t, f, B)\; dt \tag{7.17}$$

$$\hat{Q}_{xy}(f) = \frac{1}{BT}\int_0^T x(t, f, B)\; y^{o}(t, f, B) \tag{7.18}$$

where the notation ^{o} refers to a 90^{o} phase shift in one of the records. Because of the similarity of Eq. (7.17) and Eq. (7.18), it is clear that expressions for the various errors in each will be similar. Therefore, only errors in estimating $C_{xy}(f)$ will be stated below.

Due to the finite bandwidth of the filter, the estimate will be biased and that bias is given by

$$b[\hat{C}_{xy}(f)] = E[\hat{C}_{xy}(f) - C_{xy}(f)] \tag{7.19}$$

This is approximated by

$$b[\hat{C}_{xy}(f)] \approx \frac{B^2}{24} C''_{xy}(f)$$

No convenient formulas exist for the variance of the estimate; however, the variance of either estimate is bounded by

$$V[\hat{C}_{xy}(f)] \leq \frac{G_x(f)\, G_y(f)}{BT} \qquad (7.20$$

where this bound is approached as x(t) and y(t) become independent.

For the case where the center frequency of the filters are scanned continuously over the frequency interval of interest, an additional bias term is generated. Suppose that during the time interval, T, the center frequency is changed from f to f + Δf. Then the frequency scan rate is

$$\gamma = \frac{\Delta f}{T} \qquad (7.21)$$

and the cross-spectral estimate becomes

$$\hat{C}_{xy(s)}(f + \Delta f) = \frac{1}{BT} \int_0^T x(t, f + \gamma t, B)\, y(t, f + \gamma t, B)\, dt \qquad (7.22)$$

As was the case in continuous scan cross-correlation, the scan rate will be small and Eq. (7.22) can be written as

$$\hat{C}_{xy(s)}(f + \Delta f) \cong \frac{1}{BT} \int_0^T \left[x(t, f, B) + \gamma t \frac{\partial x(t, f, B)}{\partial f} \right] \left[y(t, f, B) + \gamma t \frac{\partial y(t, f, B)}{\partial f} \right] dt \tag{7.23}$$

Upon taking expected values in Eq. (7.23), it is seen that

$$E[\hat{C}_{xy(s)}(f + \Delta f)] = \hat{C}_{xy}(f) + \frac{\gamma}{BT} \int_0^T t\, x(t, f, B) \frac{\partial y(t, f, B)}{\partial f}\, dt$$

$$+ \frac{\gamma}{BT} \int_0^T t\, y(t, f, B) \frac{\partial x(t, f, B)}{\partial f}\, dt \tag{7.24}$$

$$+ \frac{\gamma^2}{BT} \int_0^T t^2 \frac{\partial x(t, f, B)}{\partial f} \frac{\partial y(t, f, B)}{\partial f}\, dt$$

It has not been possible to simplify Eq. (7.24) and express the three integrals as functions of $C_{xy}(f)$ and the impulse response functions of the filters. Further work in this area is required before a complete analysis can be accomplished.

When extraneous noise is additively present at the input to the cross-spectral analyzer, the recorded data, x(t) and y(t), are replaced by

$$x_1(t) = x(t) + n_1(t)$$

$$y_1(t) = y(t) + n_2(t) \tag{7.25}$$

The corresponding cospectral estimate becomes

$$\hat{C}_{xy(n)}(f) = \frac{1}{BT}\int_0^T x_1(t, f, B)\, y_1(t, f, B)\, dt \qquad (7.26)$$

Assuming that the noise sources have zero mean and are uncorrelated with each other and with each of the records, it follows that

$$E[\hat{C}_{xy(n)}(f)] = \frac{1}{BT}\int_0^T x(t, f, B)\, y(t, f, B)\, dt \qquad (7.27)$$

which means that the noise sources introduce no additional bias error. The effect of extraneous noise sources on the variance of the estimator is similar in nature to the effects on the cross-correlation estimate; however, the analytical expression for the noise error is too complex to be of practical value.

7.1.3 R-C Averaging

Instead of the ideal integrator which was employed in the cross-correlation and cross-spectral estimation procedures described previously, an R-C averaging circuit is frequently used. The effect of this type of averaging will be shown below.

For the cross-correlation estimate we have

$$\hat{R}_{xy(R-C)}(\tau) = \frac{1}{K}\int_0^T x(T-u)\, y(T-u+\tau)\, e^{-u/K}\, du \qquad (7.28)$$

where K = RC = time constant of RC circuit. Upon taking the expected value of both sides of Eq. (7.28) and carrying out the integration, it is seen that

$$E[\hat{R}_{xy(R-C)}(\tau)] = R_{xy}(\tau)\,[1 - e^{-T/K}] \tag{7.29}$$

This result indicates that while a bias is produced by the R-C averaging, its effect is restricted to a scale change in amplitude and the functional form of the cross-correlation is correctly estimated. As was the case in some of the previous situations, the expression for the variance of the estimate when R-C averaging is used is too complicated to be of practical use.

A similar situation exists when R-C averaging is used in cross-spectral estimation. By direct calculation, the expected value of the estimate is

$$E[\hat{C}_{xy(R-C)}(f)] = \hat{C}_{xy}(f)\,[1 - e^{-T/K}] \tag{7.30}$$

If the bias error in $C_{xy}(f)$ is small and $\exp(-\frac{T}{K})$ is small, then their products will be small compared to either one, and then Eq. (7.30) may be written as

$$E[\hat{C}_{xy(R-C)}(f)] \simeq C_{xy}(f) + \frac{B^2 C''_{xy}(f)}{24} - C_{xy}(f)\, e^{-T/K} \tag{7.31}$$

7.2 PHASE ERRORS

It is the purpose of this subsection to examine the inaccuracies resulting in cross-correlation and cross-spectral density analyses from typical phase errors in the measurement equipment. These phase errors occur in every part of the measuring system — the transducers, signal conditioners, transmission lines, telemetry units, tape recorders, A/D converters, the analyzers, and the printout devices. These phase errors can be either static or dynamic (time-varying) in character. To ensure that accurate measurements are made, these errors must be confined to acceptable bounds or compensated for in the data reduction process. The following discussions provide guidelines for evaluating the severity of different types of phase errors and for error compensation techniques.

7.2.1 Ordinary Spectral Density and Autocorrelation Measurements

In ordinary spectral density and autocorrelation analyses, the phase errors must be split into the static and dynamic components and discussed separately. The static phase errors introduce no inaccuracies into the measurement process while the dynamic phase errors do.

Static Phase Errors

For the first case, assume that the measuring circuit has a constant time delay error τ. This type of error is encountered in every instrument in the measuring chain since they all have finite bandwidths. (Of course, restrictions must be placed on the frequency bandwidth over which the constant time delay approximation is valid for each individual instrument.) In general, those instruments having the narrowest bandwidths will cause the greatest time delay. For example, assume that the response of the transducer

can be approximated by a simple mechanical oscillator. The phase factor for this transducer can be written as (Reference 38).

$$\phi(f) = \tan^{-1}\left[\frac{2\zeta\frac{f}{f_n}}{1-\left(\frac{f}{f_n}\right)^2}\right] \tag{7.32}$$

where

$\phi(f)$ = the phase factor

ζ = the damping ratio

f = the frequency in cps of the measured data

f_n = the undamped natural frequency in cps of the transducer

Equation (7.32) can be expanded in series form

$$\phi(f) = \left[\frac{2\zeta\frac{f}{f_n}}{1-\left(\frac{f}{f_n}\right)^2}\right] - \frac{1}{3}\left[\frac{2\zeta\frac{f}{f_n}}{1-\left(\frac{f}{f_n}\right)^2}\right]^3 + \frac{1}{5}\left[\frac{2\zeta\frac{f}{f_n}}{1-\left(\frac{f}{f_n}\right)^2}\right]^5 \cdots \tag{7.33}$$

If the restrictions $f \leq 0.1\, f_n$ and $\zeta \leq 1.0$ are applied, then

$$\phi(f) \approx \frac{2\zeta f}{f_n}$$

The time delay, τ_0, is a constant under these restrictions since

$$\tau_0 = \frac{\phi(f)}{2\pi f} = \frac{\zeta}{\pi f_n} \tag{7.34}$$

The autocorrelation function is defined as

$$R_x(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t + \tau)\, dt \tag{7.35}$$

Let the measured signal be

$$x^m(t) = x(t + \tau_0)$$

where τ_0 is the time delay in the measuring circuitry. The measured autocorrelation function is now

$$R_x^m(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x^m(t)\, x^m(t + \tau)\, dt \tag{7.36}$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t + \tau_0)\, x(t + \tau_0 + \tau)\, dt$$

Let $u = t + \tau_0$; $du = dt$

$$R_x^m(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T+\tau_0}^{T+\tau_0} x(u)\, x(u+\tau)\, du$$

$$= R_x(\tau)$$

Since the process being analyzed is stationary and ergodic, no error is introduced into the autocorrelation function by computing from $(-T + \tau_0)$ to $(T + \tau_0)$ instead of from $(-T)$ to (T). Because the ordinary spectral density is the Fourier transform (Wiener-Khintchine Theorem, Reference 41) of the autocorrelation function, it can be seen that the constant time delay error also introduces no error in the spectral density analysis.

Next consider the case of a phase shift that is independent of frequency.

$$\phi = k \quad ; \quad \text{where } k = \text{a constant} \tag{7.37}$$

The true spectral density, the two-sided form is used for simplicity, is

$$S_x(f) = \lim_{T\to\infty} \frac{1}{2T} \left[\int_{-T}^{T} x(t)\, e^{-j2\pi ft}\, dt\right]\left[\int_{-T}^{T} x(t)\, e^{j2\pi ft}\, dt\right] \tag{7.38}$$

where

$S_x(f)$ = the two-sided ordinary spectral density function

$x(t)$ = the data signal

Let the measured signal be

$$x^m(t) = x(t + k/2\pi f) \tag{7.39}$$

Then the measured spectral density becomes

$$S_x^m(f) = \lim_{T\to\infty} \frac{1}{2T}\left[\int_{-T}^{T} x(t + k/2\pi f)\, e^{-j2\pi ft}\, dt\right]\left[\int_{-T}^{T} x(t + k/2\pi f)\, e^{j2\pi ft}\, dt\right] \tag{7.40}$$

Let

$$\begin{aligned} u &= t + k/2\pi f \\ du &= dt \end{aligned} \tag{7.41}$$

Then

$$S_x^m(f) = \lim_{T\to\infty} \frac{1}{2T}\left[\int_{-T+k/2\pi f}^{T+k/2\pi f} x(u)\, e^{-j2\pi f(u-k/2\pi f)}\, du\right]$$

$$\left[\int_{-T+k/2\pi f}^{T+k/2\pi f} x(u)\, e^{j2\pi f(u-k/2\pi f)}\, du\right]$$

$$= \lim_{T\to\infty} \frac{\exp(jk)\exp(-jk)}{2T}\left[\int_{-T+k/2\pi f}^{T+k/2\pi f} x(u)\, e^{-j2\pi fu}\, du\right] \qquad (7.42)$$

$$\left[\int_{-T+k/2\pi f}^{T+k/2\pi f} x(u)\, e^{j2\pi fu}\, du\right]$$

$$= S_x(f)$$

Since the measured spectral density equals the true spectral density, it can also be stated through the Wiener-Khintchine Theorem that a constant phase error will not cause any error in the measured autocorrelation function.

The above result can be extended to cover the general case where the phase angle becomes some arbitrary function of frequency

$$\phi(f) = g(f) \qquad (7.43)$$

The substitution becomes

$$u = t + \frac{g(f)}{2\pi f} \quad \text{and} \quad du = dt \tag{7.44}$$

Since the operation in Eq. (7.40) is essentially a narrowband filtering operation, $g(f_0)$ is a constant in the narrow frequency band about f_0. Hence, the measured spectral density equals the true spectral density. (The fact that g(f) results in a new constant for each value of f does not influence the result since a new value of $S_x^m(f)$ is computed for each value of f.) Again, this means that there will be no error in the autocorrelation function due to a static phase error.

Dynamic Phase Errors

Dynamic phase errors occur in magnetic tape recorders from nonuniform velocity of the magnetic tape over the record and reproduce heads. Let an arbitrary function of time, h(t), represent the dynamic phase error. Then the measured data has the form

$$x^m(t) = x[t + h(t)] \tag{7.45}$$

The measured autocorrelation function is

$$\hat{R}_x^m(\tau) = \frac{1}{2T} \int_{-T}^{T} x^m(t)\, x^m(t+\tau)\, dt \tag{7.46a}$$

$$= \frac{1}{2T} \int_{-T}^{T} x[t + h(t)]\, x[t + \tau + h(t+\tau)]\, dt \tag{7.46b}$$

($\hat{R}(\tau)$ represents an estimate of $R(\tau)$, the limiting value obtained as $T \to \infty$.) Because $h(t)$ is small (0.25 to 1%), Eq. (7.46b) can be approximated by letting

$$x[t + h(t)] \approx x(t) + h(t)\, \dot{x}(t) \tag{7.47}$$

$$x[t + \tau + h(t+\tau)] \approx x(t+\tau) + h(t+\tau)\, \dot{x}(t+\tau) \tag{7.48}$$

where only the first two terms of the Taylor series expansion are used and the dot represents a time derivative ($\dot{x}(t) = d[x(t)]/dt$). (See Reference 9) Then,

$$\hat{R}_x^m(\tau) \approx \frac{1}{2T} \int_{-T}^{T} [x(t) + h(t)\, \dot{x}(t)]\, [x(t+\tau) + h(t+\tau)\, \dot{x}(t+\tau)]\, dt$$

$$\approx \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t+\tau)\, dt + \frac{1}{2T} \int_{-T}^{T} h(t+\tau)\, x(t)\, \dot{x}(t+\tau)\, dt \tag{7.49}$$

$$+ \frac{1}{2T} \int_{-T}^{T} \dot{x}(t)\, x(t+\tau)\, h(t)\, dt + \frac{1}{2T} \int_{-T}^{T} \dot{x}(t)\, \dot{x}(t+\tau)\, h(t)\, h(t+\tau)\, dt$$

If we repeat our measurement many times, we can find the expected value of the measured correlation function

$$R_x^m(\tau) = E[\hat{R}_x^m(\tau)] \tag{7.50}$$

Separating out the individual components of Eq. (7.49), one finds

$$E\left[\frac{1}{2T}\int_{-T}^{T} x(t)\, x(t+\tau)\, dt\right] = R_x(\tau) \tag{7.51}$$

and since $x(t)$ and $h(t)$ are independent and stationary

$$E\left[\frac{1}{2T}\int_{-T}^{T} x(t)\, \dot{x}(t+\tau)\, h(t+\tau)\, dt\right] = \frac{R_x'(\tau)}{2T}\int_{-T}^{T} h(t+\tau)\, dt \quad ; \text{ when } h(t+\tau) \text{ is deterministic}$$

$$= 0 \quad ; \text{ when } h(t+\tau) \text{ is random with zero mean} \tag{7.52}$$

where the slash mark denotes a derivative with respect to τ

$$\left(R'(\tau) = \frac{d[R(\tau)]}{d\tau}\right).$$

$$E\left[\frac{1}{2T}\int_{-T}^{T}\dot{x}(t)\,x(t+\tau)\,h(t)\,dt\right]=\frac{-R_x'(\tau)}{2T}\int_{-T+\tau}^{T+\tau}h(t-\tau)\,dt \quad ; \text{ when } h(t-\tau) \text{ is deterministic} \tag{7.53}$$

$$= 0 \quad ; \text{ when } h(t-\tau) \text{ is random with zero mean}$$

$$E\left[\frac{1}{2T}\int_{-T}^{T}\dot{x}(t)\,\dot{x}(t+\tau)\,h(t)\,h(t+\tau)\,dt\right]=-R_x''(\tau)\,R_h(\tau) \tag{7.54}$$

With a magnetic tape recorder, two types of dynamic phase errors occur from flutter. The first type of flutter is approximately sinusoidal and results from rotational unbalances. The second type of flutter consists of band-limited random noise. Examination of Eqs. (7.52) and (7.53) when h(t) has a sinusoidal form reveals that in taking the expected value of these equations, one would expect them to go to zero since the phase angle of the sinusoid is an unknown random variable and is uniformly distributed about zero. Thus, for the magnetic tape recorder

$$R_x^m(\tau)\approx R_x(\tau)-R_x''(\tau)\,R_h(\tau) \tag{7.55}$$

Davies (Reference 42) shows that for a sinusoidal data signal modulated by one sinusoidal flutter component when frequency modulation recording is used, the dynamic phase error is

$$h(t)=\frac{b}{2\pi f_3}\sin 2\pi f_3 t \tag{7.56}$$

where

b = the peak flutter amplitude

f_3 = the flutter frequency

Obviously, the lower the value of b and the higher the value of f_3, the less will the measured autocorrelation function vary from the true autocorrelation function.

To determine the effect of dynamic phase errors on the spectral density function, the Fourier transform of the measured autocorrelation function is taken.

$$S_x^m(f) = \int_{-\infty}^{\infty} R_x^m(\tau)\, e^{-j2\pi f\tau} d\tau$$

$$\int_{-\infty}^{\infty} [R_x(\tau) - R_x''(\tau)\, R_h(\tau)]\, e^{-j2\pi f\tau}\, d\tau \qquad (7.57)$$

$$S_x(f) - \int_{-\infty}^{\infty} R_x''(\tau)\, R_h(\tau)\, e^{-j2\pi f\tau}\, d\tau$$

Thus, the measured spectral density also has a bias error dependent upon the second derivative of the data correlation function and upon the correlation function of the dynamic phase error.

7.2.2 Cross-Correlation and Cross-Spectral Density Measurements

In the measurement of cross-correlation or cross-spectral density functions, both static and dynamic phase errors contribute to inaccuracies. In

the following paragraphs, typical types of phase errors will be examined to demonstrate the inaccuracies that they cause.

Constant Time Delay Error

The true cross-correlation function is defined as

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\, y(t+\tau)\, dt \tag{7.58}$$

Let a constant time delay error, τ_0, be introduced into y(t) relative to x(t). This could be from using transducers with different natural frequencies as one example.

$$y^m(t) = y(t+\tau_0) \tag{7.59}$$

The measured cross-correlation function becomes

$$R_{xy}^m(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\, y(t+\tau_0+\tau)\, dt$$

$$= R_{xy}(\tau+\tau_0)$$

which means that the correlation function has been shifted along the τ axis by an amount of τ_0. This error can be easily compensated if we know τ_0 accurately.

To see the effect of a constant time delay error on the cross spectrum, let us take the Fourier transform of the measured cross-correlation function

$$S_{xy}^{m}(f) = \int_{-\infty}^{\infty} R_{xy}^{m}(\tau)\, e^{-j2\pi f\tau}\, d\tau \tag{7.60}$$

$$= \int_{-\infty}^{\infty} R_{xy}(\tau + \tau_0)\, e^{-j2\pi f\tau}\, d\tau \tag{7.60}$$

Making the substitution $u = \tau + \tau_0$, $du = d\tau$, one obtains

$$S_{xy}^{m}(f) = e^{j2\pi f\tau_0} \int_{-\infty}^{\infty} R_{xy}(u)\, e^{-j2\pi fu}\, du$$

$$= e^{j2\pi f\tau_0}\, S_{xy}(f) \tag{7.61}$$

The cross spectrum is frequently computed by measuring its real part (cospectrum) and imaginary part (quadspectrum), or the magnitude and phase factor of the cross spectrum

$$S_{xy}(f) = C_{xy}(f) - jQ_{xy}(f) \tag{7.62}$$

$$= \left| S_{xy}(f) \right| e^{-j\phi_{xy}(f)} \tag{7.63}$$

where

$C_{xy}(f)$ = the cospectrum

$Q_{xy}(f)$ = the quadspectrum

$\phi_{xy}(f)$ = the phase factor

It can be shown that a constant time delay type of phase error will result in the following measured values.

$$C_{xy}^{m}(f) = C_{xy}(f)\left[\cos 2\pi f\tau_0 + (\sin 2\pi f\tau_0)\left(\frac{Q_{xy}(f)}{C_{xy}(f)}\right)\right] \tag{7.64}$$

$$Q_{xy}^{m}(f) = Q_{xy}(f)\left[\cos 2\pi f\tau_0 - (\sin 2\pi f\tau_0)\left(\frac{C_{xy}(f)}{Q_{xy}(f)}\right)\right] \tag{7.65}$$

$$\left|S_{xy}^{m}(f)\right| = \left|S_{xy}(f)\right| \tag{7.66}$$

$$\phi_{xy}^{m}(f) = \phi_{xy}(f) - 2\pi f\tau_0 \tag{7.67}$$

Constant Percentage Time Delay Error

A constant percentage time delay error can occur in the correlation analyzer if the time delay mechanism or the readout is improperly calibrated.

The measured value of the delayed signal is

$$y^{m}(t + \tau) = y(t + k\tau) \quad ; \quad \text{where } k = \text{a constant}$$

The measured cross-correlation function with this type of phase error is

$$R_{xy}^{m}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, y(t + k\tau)\, dt$$

Let

$$u = k\tau$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, y(t + u)\, dt \qquad (7.68)$$

$$= R_{xy}(u) = R_{xy}(k\tau)$$

The measured cross-spectral density function resulting when there is a constant percentage time delay type of error is

$$S_{xy}^{m}(f) = \int_{-\infty}^{\infty} R_{xy}^{m}(\tau)\, e^{-j2\pi f\tau}\, d\tau$$

$$= \int_{-\infty}^{\infty} R_{xy}(k\tau)\, e^{-j2\pi f\tau}\, d\tau$$

Let $u = k\tau$ and $du = kd\tau$

$$S_{xy}^{m}(f) = \frac{1}{k} \int_{-\infty}^{\infty} R_{xy}(u)\, e^{-j2\pi(f/k)u}\, du$$

$$= \frac{1}{k} S_{xy}(f/k) \qquad (7.69)$$

Thus, it can be seen that there is a scale factor error introduced as well as an expansion of the frequency axis.

Constant Percentage Frequency Error

Constant percentage frequency errors occur when the magnetic tape speed during reproduce differs from the recording speed or when the frequency axis of the spectral analyzer is improperly calibrated. By analogy to Eq. (7.68),

$S_{xy}^{m}(f) = S_{xy}(kf)$; where k is a constant (the frequency calibration error)

The cross-correlation function measured with a constant percentage frequency error is

$$R_{xy}^{m}(\tau) = \int_{-\infty}^{\infty} S_{xy}(kf)\, e^{j2\pi f\tau}\, df$$

$$= \frac{1}{k} R_{xy}(\tau/k) \qquad (7.70)$$

so that the constant percentage frequency type of phase error introduces a scale factor error and a tau axis expansion in the cross-correlation function analysis.

Constant Frequency Offset Error

Constant frequency offset errors can occur in heterodyne portions of the measuring circuit or in the frequency spectrum analyzer readout. For example, assume that there is a 5 Hz offset error. Then the true data at 100, 500, and 2000 Hz is plotted at 105, 505, and 2005 Hz, respectively. The measured spectral density is

$$S_{xy}^{m}(f) = S_{xy}(f - f_0) \quad ; \quad f - f_0 \geq 0$$

$$S_{xy}^{m}(f) = S_{xy}(f + f_0) \quad ; \quad f + f_0 \leq 0 \qquad (7.71)$$

$$S_{xy}^{m}(f) = 0 \quad ; \quad -f_0 < f < f_0$$

where f_0 is the frequency offset error. This error in the cross-spectral density can be easily corrected if f_0 is known accurately.

When computing the effect of a frequency offset error on cross-correlation analysis, the Fourier transform of the cross spectrum must be separated into two parts because the positive frequency portion of the cross spectrum is being translated in an opposite direction to the negative portion.

$$R^{m}_{xy}(\tau) = \int_{-\infty}^{0} S_{xy}(f + f_0)\, e^{j2\pi f\tau}\, df + \int_{0}^{\infty} S_{xy}(f - f_0)\, e^{j2\pi f\tau}\, df$$

Let

$$-(f + f_0) = v \quad ; \quad dv = -df$$

and

$$f - f_0 = u \quad ; \quad du = df$$

Then

$$R^{m}_{xy}(\tau) = -\int_{\infty}^{-f_0} S_{xy}(-v)\, e^{j2\pi(-v-f_0)\tau}\, dv + \int_{-f_0}^{\infty} S_{xy}(u)\, e^{j2\pi(u+f_0)\tau}\, du$$

$$= \int_{-f_0}^{\infty} S_{xy}(-v)\, e^{j2\pi(-v-f_0)\tau}\, dv + \int_{-f_0}^{\infty} S_{xy}(u)\, e^{j2\pi(u+f_0)\tau}\, du$$

and since

$$\int_{-f_0}^{0} S_{xy}(-v)\, e^{j2\pi(-v-f_0)\tau}\, dv = 0$$

and

$$\int_{-f_0}^{0} S_{xy}(u)\, e^{j2\pi(u+f_0)\tau}\, du = 0$$

$$R_{xy}^{m}(\tau) = e^{-j2\pi f_0\tau} \int_0^{\infty} S_{xy}(-v)\, e^{-j2\pi v\tau}\, dv + e^{j2\pi f_0\tau} \int_0^{\infty} S_{xy}(u)\, e^{j2\pi u\tau}\, du$$

$$= [\cos 2\pi f_0\tau - j \sin 2\pi f_0\tau] \int_0^{\infty} [C_{xy}(v) + jQ_{xy}(v)] [\cos 2\pi v\tau - j \sin 2\pi v\tau]\, dv$$

$$+ [\cos 2\pi f_0\tau + j \sin 2\pi f_0\tau] \int_0^{\infty} [C_{xy}(u) - jQ_{xy}(u)] [\cos 2\pi u\tau + j \sin 2\pi u\tau]\, du$$

Inspection of the above equation reveals that the two terms are complex conjugates of the form A* + A or (a - jb) + (a + jb). Therefore, the measured cross-correlation function is equal to twice the real part of either term.

$$R_{xy}^{m}(\tau) = 2\,\mathrm{Re}[A]$$

$$= [2 \cos 2\pi f_0 \tau] \int_0^\infty [C_{xy}(u) \cos 2\pi u\tau + Q_{xy}(u) \sin 2\pi u\tau]\, du$$

$$- [2 \sin 2\pi f_0 \tau] \int_0^\infty [C_{xy}(u) \sin 2\pi u\tau - Q_{xy}(u) \cos 2\pi u\tau]\, du \qquad (7.72)$$

$$= R_{xy}(\tau) \cos 2\pi f_0 \tau - [2 \sin 2\pi f_0 \tau] \int_0^\infty [C_{xy}(u) \sin 2\pi u\tau - Q_{xy}(u) \cos 2\pi u\tau]\, du$$

From Eq. (7.72) it can be seen that the effect of a constant frequency offset error on the cross-correlation analysis is to introduce a cosine multiplication of the true correlation function plus add a bias error.

Constant Phase Shift Error

Next consider the case where the measured signal y(t) is shifted by a constant number of degrees with respect to the measured signal x(t). This error most frequently occurs in the phase mismatch of cross-spectral density analyzers. From Eqs. (7.37) and (7.39),

$$y_{f,B}^{m}(t) = y_{f,B}(t + \phi_0/2\pi f)$$

where

$y_{f,B}(t)$ = the data signal passed through a bandpass filter of bandwidth B and center frequency f.

The cospectrum can be defined as

$$C_{xy}(f) = \lim_{\substack{T\to\infty \\ B\to 0}} \frac{1}{2BT} \int_{-T}^{T} x_{f,B}(t)\, y_{f,B}(t)\, dt \tag{7.73}$$

and similarly, the quadspectrum can be defined as

$$Q_{xy}(f) = \lim_{\substack{T\to\infty \\ B\to 0}} \frac{1}{2BT} \int_{-T}^{T} x_{f,B}(t)\, y_{f,B}\left(t + \frac{\pi/2}{2\pi f}\right) dt \tag{7.74}$$

The measured cospectrum becomes

$$C_{xy}^{m}(f) = \lim_{\substack{T\to\infty \\ B\to 0}} \frac{1}{2BT} \int_{-T}^{T} x(t)\, y\left(t + \frac{\phi_0}{2\pi f}\right) dt$$

Removing the bandwidth limiting process, an estimate of the cospectrum is obtained.

$$C_{xy}^{m}(f) = \lim_{T\to\infty} \frac{1}{2BT} \int_{-T}^{T} x(t)\, y\left(t + \frac{\phi_0}{2\pi f}\right) dt$$

$$= \frac{1}{B}\left[R_{xy,f,B}\left(\frac{\phi_0}{2\pi f}\right)\right] \tag{7.75}$$

where $R_{xy,\,f,\,B}\left(\frac{\phi_0}{2\pi f}\right)$ = the cross-correlation function, evaluated at $\tau = \phi_0/2\pi f$, obtained when both x(t) and y(t) are passed through narrow bandpass filters of bandwidth B and center frequency f.

Similarly, an estimate of the measured quadspectrum is obtained.

$$\hat{Q}^{m}_{xy}(f) = \frac{1}{B}\left[R_{xy,\,f,\,B}\left(\frac{\phi_0 + \frac{\pi}{2}}{2\pi f}\right)\right] \tag{7.76}$$

The true phase factor is

$$\phi_{xy}(f) = \tan^{-1}\frac{Q_{xy}(f)}{C_{xy}(f)}$$

and the measured phase factor is

$$\phi^{m}_{xy}(f) = \tan^{-1}\frac{Q^{m}_{xy}(f)}{C^{m}_{xy}(f)} = \tan^{-1}\left\{\frac{\frac{1}{B}\left[R_{xy,\,f,\,B}\left(\frac{\phi_0 + \frac{\pi}{2}}{2\pi f}\right)\right]}{\frac{1}{B}\left[R_{xy,\,f,\,B}\left(\frac{\phi_0}{2\pi f}\right)\right]}\right\}$$

$$= \tan^{-1}\left\{\frac{R_{xy,\,f,\,B}\left(\frac{\phi_0 + \frac{\pi}{2}}{2\pi f}\right)}{R_{xy,\,f,\,B}\left(\frac{\phi_0}{2\pi f}\right)}\right\} \tag{7.77}$$

To evaluate Eq. (7.77), it is necessary to assume some specific form for the correlation function. Assume that single-tuned filters are used to compute the cross-spectrum and to simplify the calculations assume that $y(t) = x(t)$. Then, (from Reference 9)

$$R_{xy,\,f,\,B}(\tau) = R_{x,\,f,\,B}(\tau) = R_x(0)\, e^{-b|\tau|} \cos 2\pi f_0 \tau \qquad (7.78)$$

where

$2b$ = the half power bandwidth of the filter

f_0 = the center frequency of the filter

$$\phi_{xy}^{m}(f_0) = \tan^{-1}\left\{\frac{e^{-b\left|(\phi_0+\pi/2)/2\pi f_0\right|}\cos\left[(2\pi f_0)\left(\frac{\phi_0+\frac{\pi}{2}}{2\pi f_0}\right)\right]}{e^{-b\left|\phi_0/2\pi f_0\right|}\cos\left[(2\pi f_0)\left(\frac{\phi_0}{2\pi f_0}\right)\right]}\right\}$$

$$= \tan^{-1}\left\{\left[e^{-b/\left|(1/4 f_0\right|}\right]\left[-\tan\phi_0\right]\right\}$$

The "q" of the filter is $f_0/2b$, hence

$$\phi_{xy}^{m}(f_0) = \tan^{-1}\left\{\left[\left[e^{-\left|1/8q\right|}\right]\left[\tan\left(-\phi_0\right)\right]\right]\right\}$$

and if $q > 1$, as is the normal case

$$\phi_{xy}^{m}(f_0) = -\phi_0$$

As is expected, the phase factor is in error by an amount ϕ_0. [$\phi_{xy}(f_0)$ should have been zero since we assumed $y(t) = x(t)$.] Now the effect of a constant phase error on the magnitude of the cross-spectral density will be determined.

$$\left|S_{xy}^{m}(f)\right| = \left\{\left[C_{xy,\,i,\,B}^{m}(f)\right]^2 + \left[Q_{xy,\,f,\,B}^{m}(f)\right]^2\right\}^{\frac{1}{2}}$$

$$\left|S_{xy}^{m}(f)\right| = \left\{\left[\frac{1}{B}R_{xy,\,f,\,B}\left(\frac{\phi_0}{2\pi f}\right)\right]^2 + \left[\frac{1}{B}R_{xy,\,f,\,B}\left(\frac{\phi_0 + \frac{\pi}{2}}{2\pi f}\right)\right]^2\right\}^{\frac{1}{2}}$$

Assume, as before, that $x(t) = y(t)$ and the analysis filters are of the single-tuned variety.

$$\left| S^{m}_{xy}(f) \right| = \frac{R_{x,f,B}(0)}{B} \left\{ \left[e^{-b\left|\phi_0/2\pi f\right|} \cos 2\pi f \left(\frac{\phi_0}{2\pi f} \right) \right]^2 + \left[e^{-b\left|\frac{\phi_0 + \frac{\pi}{2}}{2\pi f}\right|} \cos 2\pi f \left(\frac{\phi_0 + \frac{\pi}{2}}{2\pi f} \right) \right]^2 \right\}^{\frac{1}{2}}$$

$$= \frac{R_{x,f,B}(0)}{B} \left\{ \left[e^{-\left|\phi_0/4\pi q\right|} \cos \phi_0 \right]^2 + \left[e^{-\left|\phi_0 + \frac{\pi}{2}/4\pi q\right|} \cos \left(\phi_0 + \frac{\pi}{2}\right) \right]^2 \right\}^{\frac{1}{2}}$$

If $\phi_0 < \frac{\pi}{2}$ and $q > 1$,

$$\left| S^{m}_{xy}(f) \right| = \left\{ \frac{R^2_{x,f,B}(0)}{B^2} \left[\cos^2 \phi_0 + \sin^2 \phi_0 \right] \right\}^{\frac{1}{2}}$$

$$= \frac{R_{x,f,B}(0)}{B}$$

Therefore,

$$\left| S^{m}_{xy}(f) \right| = \left| S_{xy}(f) \right| \tag{7.79}$$

Thus, as one would intuitively expect, a constant phase error does not cause an inaccuracy in the magnitude of the cross-spectrum. The measured co and quadspectra are:

$$C_{xy}^{m}(f) = \left| S_{xy}^{m}(f) \right| \cos \phi_{xy}^{m}(f)$$

$$= C_{xy}(f) \left[\frac{\cos \phi_{xy}^{m}(f)}{\cos \phi_{xy}(f)} \right] \tag{7.80}$$

$$Q_{xy}^{m}(f) = \left| S_{xy}^{m}(f) \right| \sin \phi_{xy}^{m}(f)$$

$$= Q_{xy}(f) \left[\frac{\sin \phi_{xy}^{m}(f)}{\sin \phi_{xy}(f)} \right] \tag{7.81}$$

The inaccuracy in the cross-correlation function can be found by transforming the cross spectrum. The transform must be divided into two parts at f equal to zero since the constant phase error causes the phase factor for positive frequencies to be translated in the opposite direction to the phase factor for negative frequencies. This can be seen from the symmetry properties of the cross spectrum. (See Reference 11)

$$C_{xy}(-f) = C_{xy}(f) \tag{7.82}$$

$$Q_{xy}(-f) = -Q_{xy}(f) \tag{7.83}$$

Since

$$\phi(f) = \tan^{-1} \frac{Q_{xy}(f)}{C_{xy}(f)}$$

$$\phi_{xy}(-f) = -\phi_{xy}(f) \tag{7.84}$$

Let the measured phase factor at positive frequencies be

$$\phi^{m}_{xy}(f) = \phi_{xy}(f) + \phi_0 \tag{7.85}$$

Then the phase factor at negative frequencies is

$$\phi^{m}_{xy}(f) = \phi_{xy}(f) - \phi_0 \tag{7.86}$$

and the cross-spectrum can be written

$$S^{m}_{xy}(f) = \left| S_{xy}(f) \right| e^{-j[\phi_{xy}(f)+\phi_0]} \quad ; \quad f \geq 0 \tag{7.87a}$$

$$= \left| S_{xy}(f) \right| e^{-j[\phi_{xy}(f)-\phi_0]} \quad ; \quad f \leq 0 \tag{7.87b}$$

or

$$S_{xy}^{m}(f) = e^{-j\phi_0} S_{xy}(f) \quad ; \quad f \geq 0 \tag{7.88a}$$

$$= e^{j\phi_0} S_{xy}(f) \quad ; \quad f \leq 0 \tag{7.88b}$$

The measured cross-correlation function is written as follows:

$$R_{xy}^{m}(\tau) = e^{-j\phi_0} \int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau}\, df + e^{j\phi_0} \int_{-\infty}^{0} S_{xy}(f)\, e^{j2\pi f\tau}\, df$$

Let $u = -f$ in the second integral and $du = -df$, then

$$R_{xy}^{m}(\tau) = e^{-j\phi_0} \int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau}\, dt - e^{j\phi_0} \int_{\infty}^{0} S_{xy}(-u)\, e^{-j2\pi u\tau}\, du$$

$$= e^{-j\phi_0} \int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau}\, df + e^{j\phi_0} \int_0^{\infty} S_{xy}(-u)\, e^{-j2\pi u\tau}\, du$$

$$= A + A^*$$

where A^* = the complex conjugate of A.

Therefore

$$R_{xy}^{m}(\tau) = 2\,\mathrm{Re}\left[e^{-j\phi_0}\int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau}\, df\right]$$

$$R_{xy}^{m}(\tau) = 2\,\mathrm{Re}\left\{(\cos\phi_0 - j\sin\phi_0)\int_0^{\infty}\left[C_{xy}(f) - jQ_{xy}(f)\right](\cos 2\pi f\tau + j\sin 2\pi f\tau)\, dt\right\} \tag{7.89}$$

$$= \cos\phi_0\, R_{xy}(\tau) + 2\sin\phi_0\int_0^{\infty}\left[C_{xy}(f)\sin 2\pi f\tau - Q_{xy}(f)\cos 2\pi f\tau\right] df$$

It can be seen that the constant phase error causes the measured cross-correlation function to be equal to the true value multiplied by a cosine term plus a bias error. Notice that this result is very similar to that in Eq. (7.72).

Frequency Dependent Phase Errors

A general expression can be obtained for the inaccuracies caused by frequency dependent phase errors. For cross-spectral measurements

$$S_{xy}^{m}(f) = S_{xy}(f)\, e^{-j\phi_0(f)} \tag{7.90}$$

where $\phi_0(f)$ = the frequency dependent phase error.

For the cross-correlation measurements

$$R_{xy}^{m}(\tau) = \int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau} \cdot e^{-j\phi_0(f)}\, df + \int_{-\infty}^{0} S_{xy}(f)\, e^{j2\pi f\tau}\, e^{-j\phi_0(f)}\, df$$

Let $u = -f$, $du = -df$ in the second integral

$$R_{xy}^{m}(f) = \int_0^{\infty} S_{xy}(f)\, e^{j2\pi f\tau} \cdot e^{-j\phi_0(f)}\, df - \int_{\infty}^{0} S_{xy}(-u)\, e^{-j2\pi u\tau} \cdot e^{-j\phi_0(-u)}\, du$$

$$= \int_0^{\infty} S_{xy}(f)\, e^{j[2\pi f\tau - \phi_0(f)]}\, df + \int_0^{\infty} S_{xy}(-u)\, e^{-j[2\pi u\tau - \phi_0(u)]}\, du$$

$$= A + A^*$$

$$= 2\,RE\left[\int_0^{\infty} S_{xy}(f)\, e^{j[2\pi f\tau - \phi_0(f)]}\, df\right] \qquad (7.91)$$

As an example, consider the case where the y(t) signal is passed through an RC lowpass filter while the x(t) signal is not.

$$x^m(t) = x(t)$$

$$y^m(t) = y(t)\ A(f)\ e^{-j\phi(f)}$$

$$A(f) = \frac{1}{\sqrt{1 + (2\pi f RC)^2}} \tag{7.92}$$

$$\phi_0(f) = \tan^{-1}\ [2\pi f RC] \tag{7.93}$$

From Eq. (7.90),

$$S^m_{xy}(\tau) \approx S_{xy}(f)\ e^{-j[\tan^{-1}(2\pi f RC)]} \tag{7.94}$$

if f is restricted to a range where $A(f) \approx 1$.

To compute the inaccuracy in the cross-correlation function, let A(f) be restricted as follows.

$$1.0 \geq A(f) \geq 0.98$$

From Eq. (7.92), we find the highest permissible frequency, f_h,

$$f_h = \frac{0.2}{2\pi RC}$$

and from Eq. (7.93), the maximum phase shift is found to be

$$\theta_{max} = \theta(f_h) = -0.2$$

From Eq. (7.91),

$$R_{xy}^{m}(\tau) = 2\,\mathrm{Re}\left[\int_0^{\infty} S_{xy}(f)\, e^{j[2\pi f\tau - \tan^{-1}(2\pi f\,RC)]}\, df\right]$$

The arc tan can be expanded as follows.

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \cdots$$

and since for the highest frequency of operation

$$\tan^{-1}(0.2) = 0.2 - \frac{(0.2)^3}{3} + \frac{(0.2)^5}{5} - \frac{(0.2)^7}{7} \cdots$$

$$\approx 0.2$$

We can write

$$-\tan^{-1}(2\pi f\,RC) \approx -2\pi f\,RC$$

Therefore,

$$R_{xy}^{m}(\tau) \approx 2\,\mathrm{Re}\left[\int_{0}^{\infty} S_{xy}(f)\, e^{j2\pi f(\tau - RC)}\, df\right]$$

$$\approx R_{xy}(\tau - RC)$$

which is the same result as for a constant time delay type of phase error. This is to be expected since the maximum frequency was restricted to a range where $\theta(f)$ is linear.

Dynamic Phase Errors

The most severe dynamic phase errors in the measurement system occur in the magnetic tape recording process. These errors are caused by the nonuniform velocity of the tape and motion of one tape track relative to another. Dynamic skew, differential flutter, and differential head stack vibration are sources of this relative motion. (See Reference 39.) This error is commonly called dynamic Interchannel Time Delay Error (ITDE). The measured values of x(t) and y(t) are

$$x^{m}(t) = x[t + g(t)] \tag{7.95}$$

$$y^{m}(t) = y[t + h(t)] \tag{7.96}$$

The estimate of the measured cross-correlation function then becomes

$$\hat{R}^{m}_{xy}(\tau) = \frac{1}{2T}\int_{-T}^{T} x[t + g(t)]\ y[t + \tau + h(t + \tau)]\ dt \qquad (7.97)$$

Following the development for the effects of dynamic phase errors on auto-correlation analyses, $x[t + g(t)]$ and $y[t + \tau + h(t + \tau)]$ are approximated by the first two terms of a Taylor series.

$$x[t + g(t)] \approx x(t) + g(t)\ \dot{x}(t)$$

$$y[t + \tau + h(t + \tau)] \approx y(t + \tau) + h(t + \tau)\ \dot{y}(t + \tau)$$

$$\hat{R}^{m}_{xy}(\tau) \approx \frac{1}{2T}\int_{-T}^{T} [x(t) + g(t)\ \dot{x}(t)]\ [y(t + \tau) + h(t + \tau)\ \dot{y}(t + \tau)]\ dt$$

$$\approx \frac{1}{2T}\int_{-T}^{T} [x(t)]\ [y(t + \tau)]\ dt + \frac{1}{2T}\int_{-T}^{T} g(t)\ \dot{x}(t)\ y(t + \tau)\ dt$$

$$+ \frac{1}{2T}\int_{-T}^{T} h(t + \tau)\ x(t)\ \dot{y}(t + \tau)\ dt + \frac{1}{2T}\int_{-T}^{T} g(t)\ h(t + \tau)\ \dot{x}(t)\ \dot{y}(t + \tau)\ dt$$

Since x(t) and y(t) are independent of g(t) and h(t), and all four are stationary

$$E[\hat{R}^{m}_{xy}(\tau)] \approx R_{xy}(\tau) - R''_{xy}(\tau)\, R_{gh}(\tau) \tag{7.98}$$

(As shown previously, the middle two terms vanish for the dynamic phase error functions typical of magnetic tape recorders.) Thus, the measured cross-correlation function is equal to the true cross-correlation function plus a bias term. This bias term is the product of the second derivative of the cross-correlation function of the data and the cross-correlation function of the dynamic phase error signals.

The effect of these dynamic phase errors on cross-spectral density analyses can be determined by taking the Fourier transform of the measured cross-correlation function.

$$S^{m}_{xy}(f) = \int_{-\infty}^{\infty} R^{m}_{xy}(\tau)\, e^{-j2\pi f\tau}\, d\tau$$

$$\approx \int_{-\infty}^{\infty} [R_{xy}(\tau) - R''_{xy}(\tau)\, R_{gh}(\tau)]\, e^{-j2\pi f\tau}\, d\tau$$

$$\approx S_{xy}(f) - \int_{-\infty}^{\infty} R''_{xy}(\tau)\, R_{gh}(\tau)\, e^{-j2\pi f\tau}\, d\tau \tag{7.99}$$

This measured cross-spectrum is then equal to the true cross-spectrum plus a bias error that also depends on the second derivative of the cross-correlation function of the data and the cross-correlation function of the dynamic phase errors.

7.3 CORRECTIONS FOR FINITE SIZE TRANSDUCERS IN THE MEASUREMENT OF BOUNDARY LAYER PRESSURE FLUCTUATIONS

Consider two transducers used to measure pressure as shown in Figure 7-1.

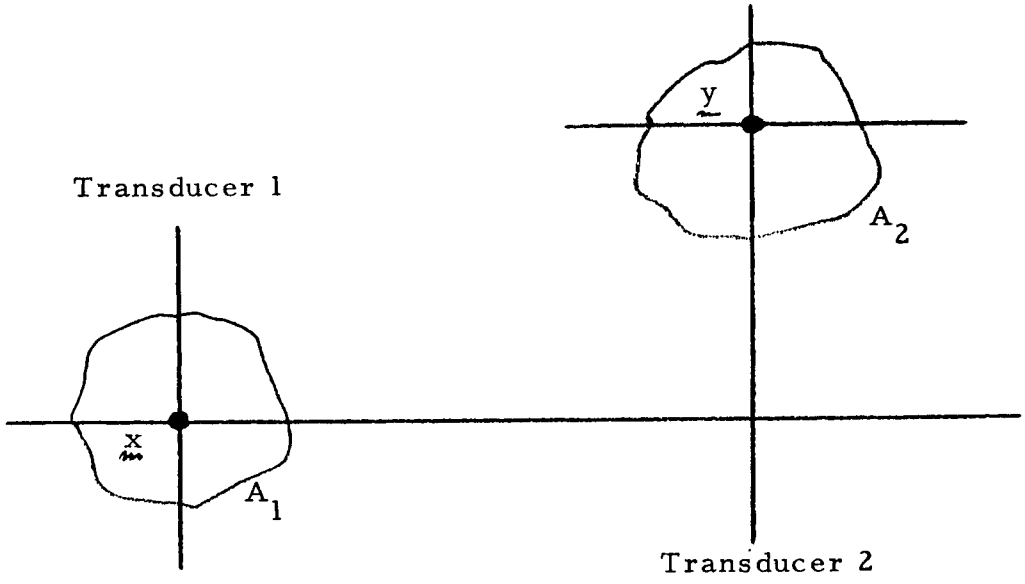


Figure 7-1. Transducers

$\underset{\sim}{x}$, $\underset{\sim}{y}$ } position vectors of a point on the transducer surface

Let $K_1(\underset{\sim}{x})$, $K_2(\underset{\sim}{y})$ be the output voltage from the transducer due to a unit point force applied at some point on the transducer surface. Then the total output voltage due to some pressure field is

$$E_1(t) = \int_{A_1} p_1(\underset{\sim}{x}, t)\, K_1(\underset{\sim}{x})\, dA_1 \tag{7.100}$$

$$E_2(t) = \int_{A_2} p_2(\underset{\sim}{y}, t)\, K_2(\underset{\sim}{y})\, dA_2 \tag{7.101}$$

Consider the cross-correlation of these two signals (or their cross power spectrum)

$$R_E^{1,2}(\tau) = \langle E_1(t)\, E_2(t+\tau) \rangle \tag{7.102}$$

where $\langle\ \rangle$ means expected value

$$R_E^{1,2}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \iint p_1(\underset{\sim}{x}, t)\, p_2(\underset{\sim}{y}, t+\tau)\, K_1(\underset{\sim}{x})\, K_2(\underset{\sim}{y})\, dA_1\, dA_2\, dt \tag{7.103}$$

$K(x)$ is determined experimentally or theoretically by determining the voltage for a point force (a needle point) and normalizing it appropriately so that $\int K(\underset{\sim}{x})\, dA = G$, when G is the output for a uniform pressure.

Only the pressure terms enter into the time averaging, so let us perform that independently

$$\langle p_1(\underset{\sim}{x}, t)\, p_2(\underset{\sim}{y}, t+\tau)\rangle = R_p(\underset{\sim}{x}, \underset{\sim}{y}, \tau) \tag{7.104}$$

This may be normalized by dividing by $\left[\overline{p^2}(\underset{\sim}{x})\,\overline{p^2}(\underset{\sim}{y})\right]^{\frac{1}{2}}$ to give the normalized cross-correlation,

$$\tilde{C}_p(\underset{\sim}{x}, \underset{\sim}{y}, \tau) = \frac{R_p(\underset{\sim}{x}, \underset{\sim}{y}, \tau)}{\left[\overline{p^2}(\underset{\sim}{x})\,\overline{p^2}(\underset{\sim}{y})\right]^{\frac{1}{2}}} \tag{7.105}$$

where

$\overline{p^2}(\underset{\sim}{x})$ = mean square pressure at $\underset{\sim}{x}$

$\overline{p^2}(\underset{\sim}{y})$ = mean square pressure at $\underset{\sim}{y}$

In a similar fashion the normalized cross-spectral density, $\tilde{C}_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega)$ is given by

$$\tilde{C}_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega) = \frac{S_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega)}{\left[S_p(\underset{\sim}{x}, \omega)\, S_p(\underset{\sim}{y}, \omega)\right]^{\frac{1}{2}}} \tag{7.106}$$

where

$S_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega)$ = cross-spectral density of the pressure at $\underset{\sim}{x}$ and $\underset{\sim}{y}$

$S_p(\underset{\sim}{x}, \omega)$ = ordinary spectral density of the pressure at $\underset{\sim}{x}$

$S_p(\underset{\sim}{y}, \omega)$ = ordinary spectral density of the pressure at $\underset{\sim}{y}$

The measured cross-spectral density, $S^{1,2}_{p,m}(\omega)$ and cross-correlation $R^{1,2}_{p,m}(\tau)$ are given by

$$S^{1,2}_{p,m}(\omega) = \iint S_p(\underset{\sim}{x}, \underset{\sim}{y}, \tau)\, \frac{K_1(\underset{\sim}{x})\, K_2(\underset{\sim}{y})}{G^2}\, dA_1\, dA_2 \tag{7.107}$$

$$R^{1,2}_{p,m}(\tau) = \iint R_p(\underset{\sim}{x}, \underset{\sim}{y}, \tau) \frac{K_1(\underset{\sim}{x})\, K_2(\underset{\sim}{y})}{G^2}\, dA_1\, dA_2 \tag{7.108}$$

where G is the sensitivity of the transducer to a uniform pressure. For a homogeneous pressure field, the cross-spectrum and cross-correlation is dependent only upon the separation of the points. Because the object of the measurement is to obtain an estimate of the correlation or spectrum at some point or points and not over some finite area, the measured function is usually taken to be that at points $\underset{\sim}{x}_0$ and $\underset{\sim}{y}_0$ (two points, perhaps the centroids, of the respective areas. Considering only the cross-spectrum now, the degradation due to a finite size transducer is

$$\frac{S^{1,2}_{p,m}(\omega)}{S_p(\underset{\sim}{x}_0, \underset{\sim}{y}_0, \omega)} = \iint \frac{S_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega)}{S_p(\underset{\sim}{x}_0, \underset{\sim}{y}_0, \omega)} \frac{K_1(\underset{\sim}{x})\, K_2(\underset{\sim}{y})}{G^2}\, dA_1\, dA_2 \tag{7.109}$$

Equation (7.109) may be evaluated analytically if the forms are known and integrable; otherwise, it may be performed by numerical integration. The degradation of resolution of the one point power spectrum is

$$\frac{S_{p,m}(\omega)}{S_p(\omega)} = \iint \tilde{C}_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega) \frac{K(\underset{\sim}{x})\, K(\underset{\sim}{y})}{G^2}\, dA_1\, dA_2 \tag{7.110}$$

where $\tilde{C}_p(\underset{\sim}{x}, \underset{\sim}{y}, \omega)$ is the cross-spectrum normalized at the reference point $\underset{\sim}{x}_0$, and $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are different positions on the same transducer face.

Corcos (Reference 43) has evaluated Eq. (7.110) for a specific boundary layer cross-spectrum function $\tilde{C}$, and for round and square transducers of uniform sensitivity (K = 1 over the area). Gilchrist and Strawderman (Reference 44) followed this analysis with work which took into account the variable sensitivity, but used a questionable approximation of the transducer size. Willmarth and Roos (Reference 45) looked at the same problem in a slightly different manner and obtained analytical results which showed Corcos to be overly conservative, however they also used a uniformly sensitive transducer.

Additional work by Chandiramani (Reference 46) has shown that the correction is a function of both the wave number spectrum and the boundary layer thickness. He examined several models of boundary layer turbulence and determined a correction curve for each model. These are shown in Figures 7-2 and 7-3. Note that Corcos' estimate of the correction appears to be overly conservative in all cases. Furthermore, the correction factor is dependent upon the ratio of transducer diameter to boundary layer thickness, a smaller ratio giving less attenuation.

Because none of the models studied by the earlier investigators took into account the variable sensitivity of the transducer element, all the correction estimates are conservative. To obtain a correction estimate which takes into account this factor, an integration of Eq. (7.110) was made with the normalized cross-spectral density in the form of

$$\tilde{C}_{p}(\underset{\sim}{x}, \underset{\sim}{y}, \omega) = e^{-C_1 \frac{\omega}{U_c} |x_1 - y_1| - C_2 \frac{\omega}{U_c} |x_2 - y_2| - j \frac{\omega}{U_c} (x_1 - y_1)} \qquad (7.111)$$

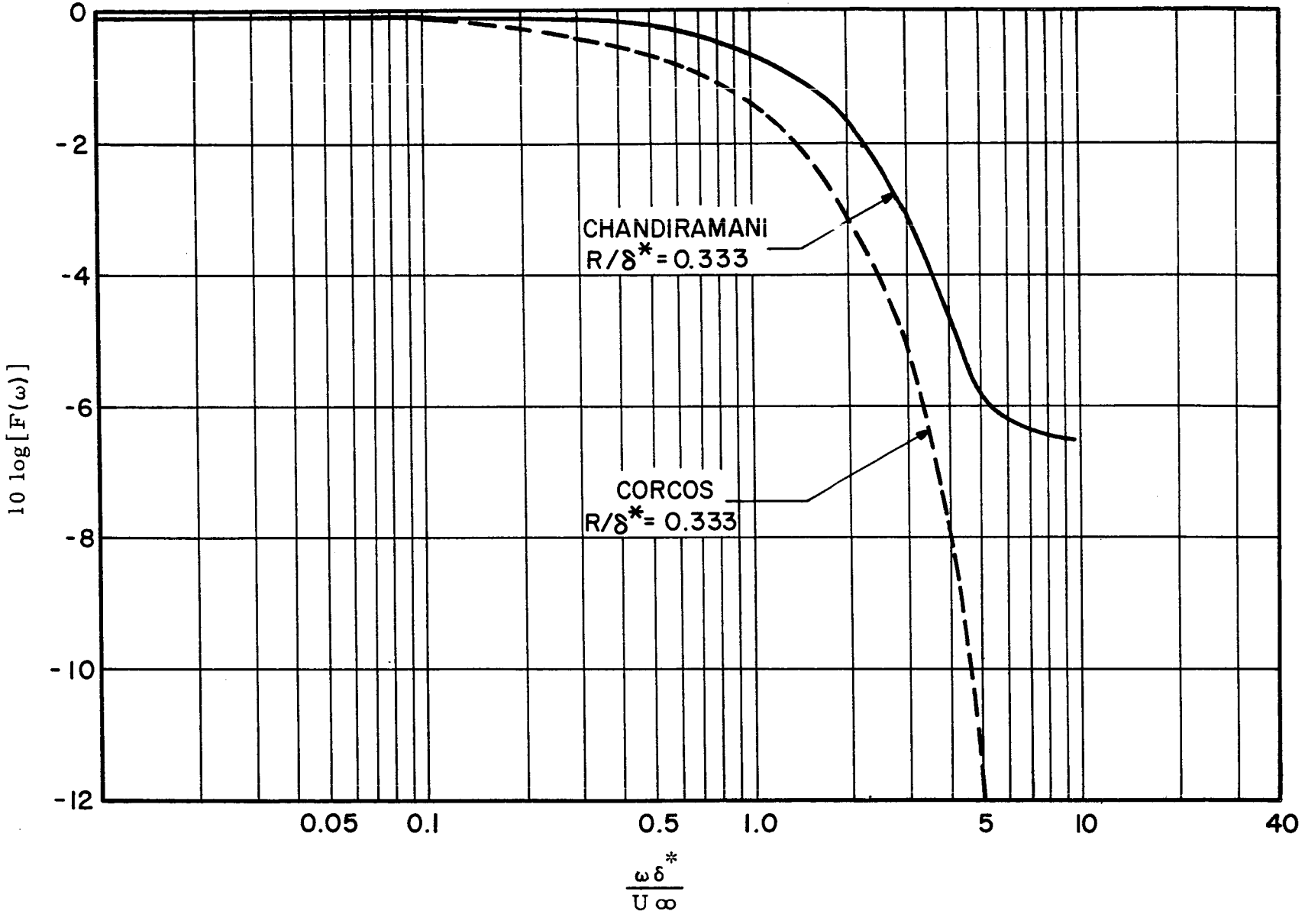


Figure 7-2. Attenuation Factor $F(\omega)$ for Circular Transducer of Uniform Sensitivity

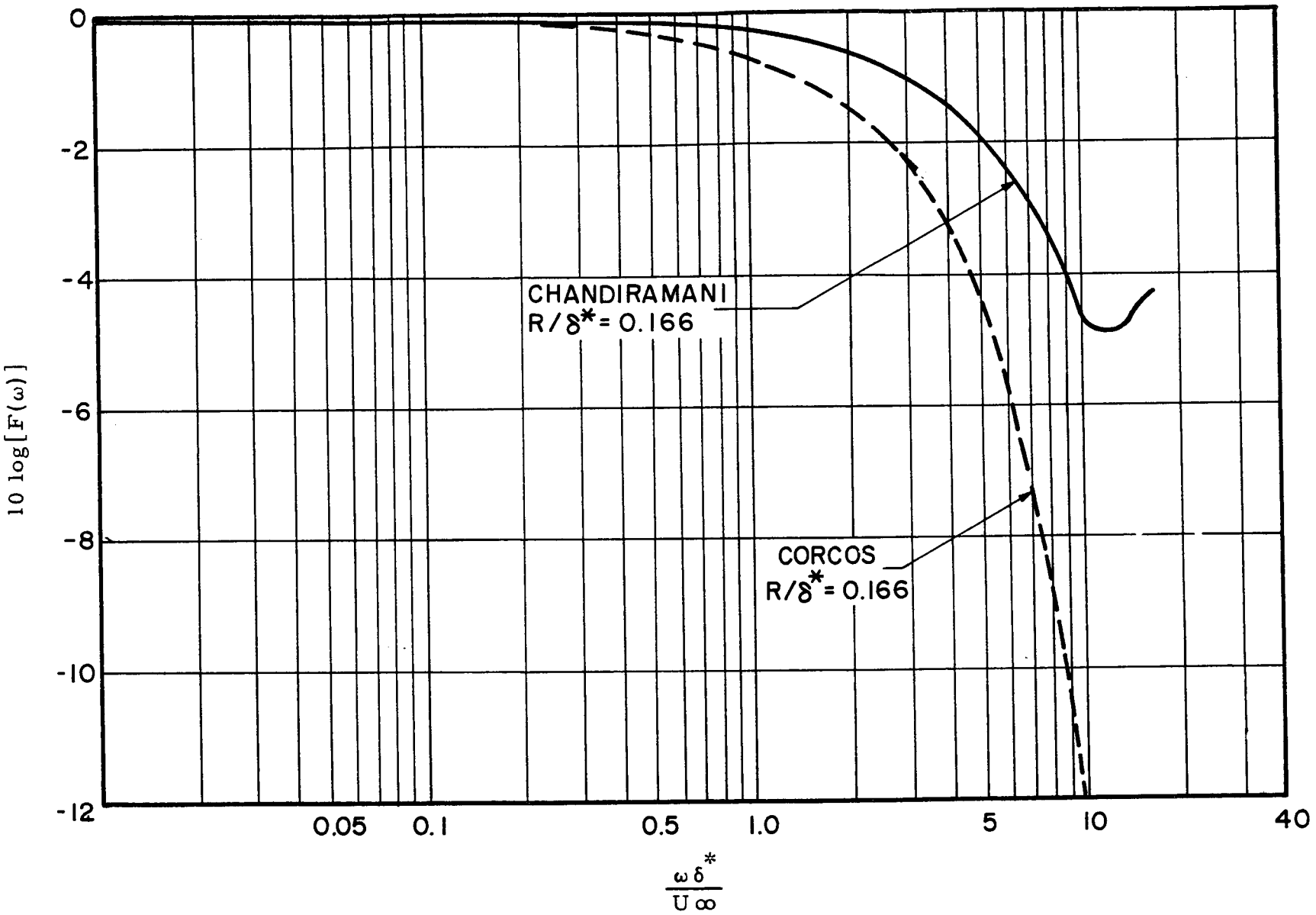


Figure 7-3. Attenuation Factor F(ω) for Circular Transducer of Uniform Sensitivity

and K(x) taken from the data of Reference 42. The sensitivity of this transducer is shown in Figure 7-4. The results are presented in Figure 7-5, and show a large change from the prediction of Corcos. The effect of the variable sensitivity is to make the effective size of the transducer smaller. Thus, if the correction of Corcos had been applied to the data, the corrected spectrum would be much too high at the high frequencies.

It is also evident from these curves that the correction factor is dependent upon the boundary layer parameters. From these facts we can conclude that the true correction factor for determining the power spectral density of the pressure is dependent upon the cross-spectral density of the pressure field and the sensitivity of the transducer. Unless both functions are accurately and reliably known, any estimate is in error. Unfortunately, the present state-of-the-art does not include this knowledge. As a result, all correction factors must be recognized as having some inherent and unknown error.

Based upon all computations to date, the model of Corcos appears to be the most conservative. Taking into account the transducer sensitivity gives an improved but still conservative estimate. So long as $r/\delta^* < 0.5$, where r is the transducer radius and δ^* is the boundary layer displacement thickness, this estimate appears to be the most practical to use. Naturally, as more experimental data becomes available and a better description of the boundary layer is possible, an improved estimate can be formulated. Until that time, the curve of Figure 7-5 should be used to correct the measured pressure power spectrum.

An additional factor which must be accounted for is the change in transducer sensitivity with frequency. As is shown in Figure 7-6, the sensitivity across the face of a condenser microphone changes drastically with frequency. While no measured data is available for other transducers, it must be assumed that there is also some frequency dependence with them also. These changes will be reflected by the characteristic curves of $S_{p,m}(\omega)/S_p(\omega)$, which may move either up or down.

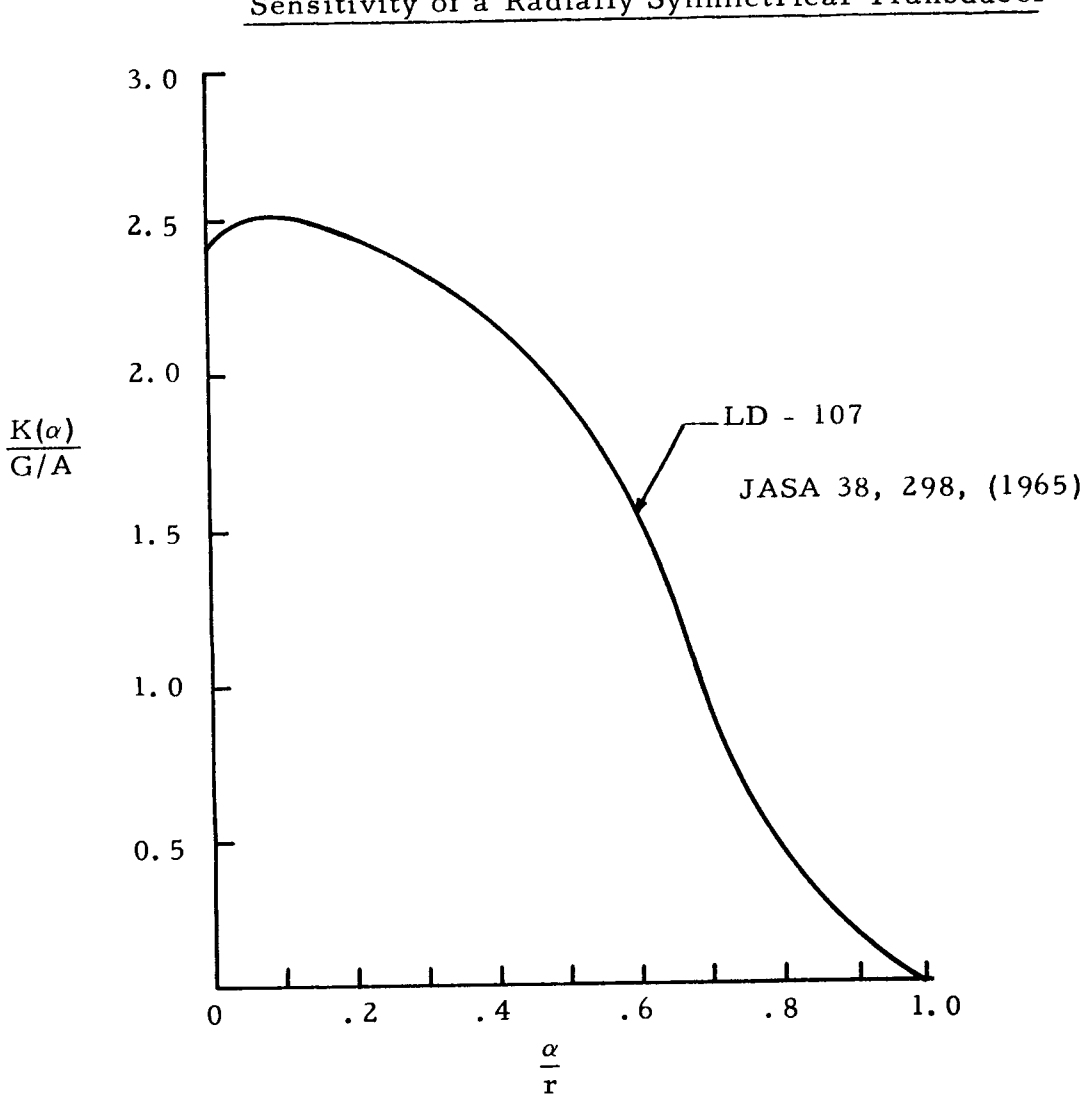


Figure 7-4. Transducer Sensitivity versus Distance from Center (α)

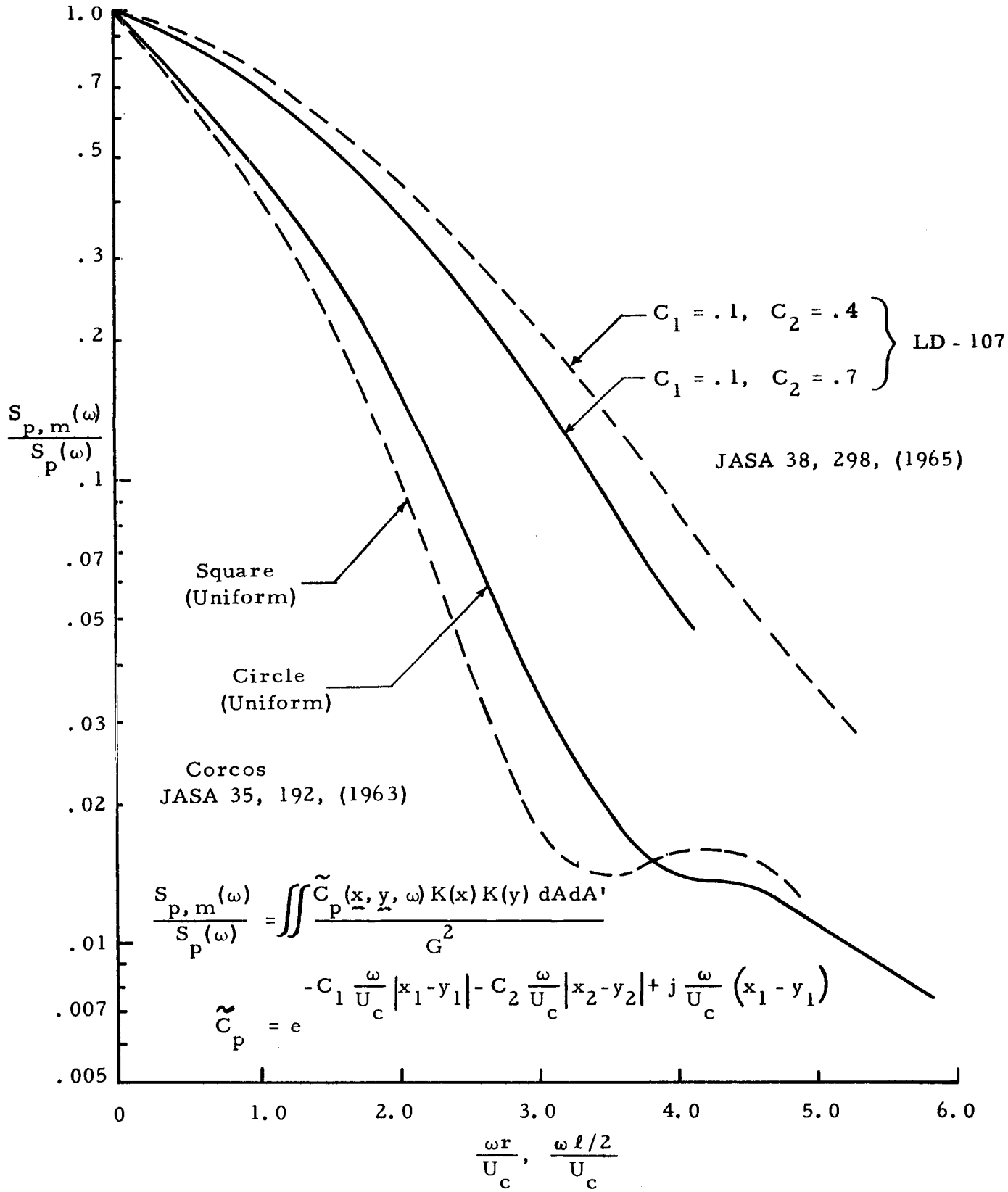


Figure 7-5. Correction Factor for Finite Size Transducer

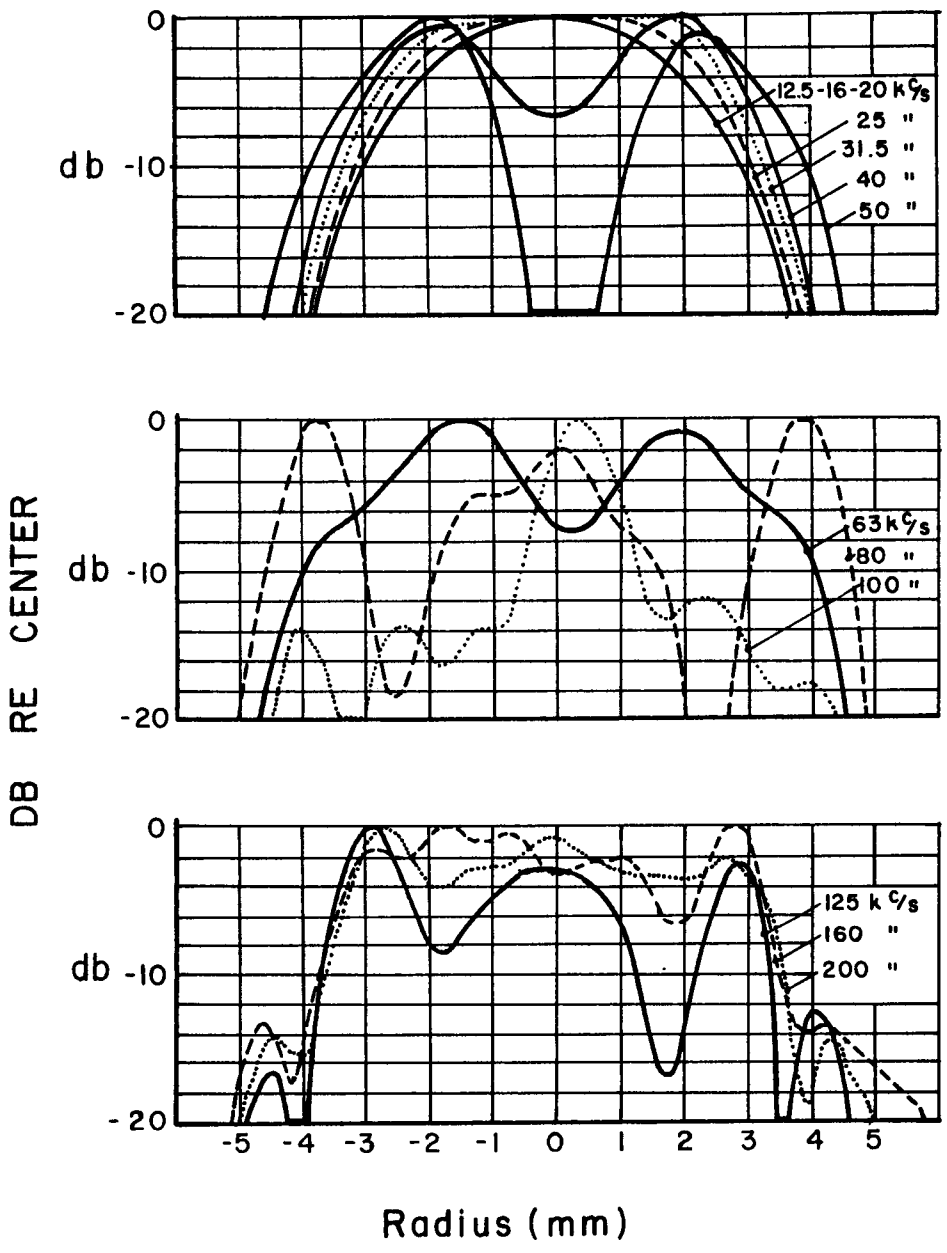


Figure 7-6. Sensitivity of B & K 4134 Microphone

7.4 MAGNETIC TAPE RECORDER DYNAMIC PHASE ERRORS

Previous reports (References 39 and 47 through 50) have considered the errors introduced into correlation and spectral density measurements by non-uniform motion of the magnetic tape used to record the data. References 39, 47, and 48 treat these errors from purely dimensional considerations. Reference 47 is a general summary of timing errors while References 39 and 49 contain discussions of phase errors between tape tracks caused by Interchannel Time Displacement Errors (ITDE). Reference 49 discusses the errors introduced into ordinary spectral density and autocorrelation functions from flutter of a single tape track. Examples are worked for both a single sinusoid flutter component and a single band-limited white random noise flutter component modulation of a single sinusoidal data signal. This report attacks the flutter problem by examining a Bessel function expansion of the frequency modulated data. (Flutter causes the frequency modulation.) Reference 50 also considers the error introduced into ordinary spectral density analyses by a single band-limited white noise flutter component modulation of a single sinusoidal data component.

A new model for studying the errors introduced into both ordinary and cross spectral and/or correlation measurements is presented in Section 7.2 of this report. Because of the importance of the dynamic phase errors introduced by recording the data on magnetic tape, this subsection examines the problem in greater detail.

7.4.1 Tape Motion Equations

Following a development similar to that used in Reference [42] let an arbitrary signal, $x(t)$ be recorded on a magnetic tape moving at a velocity $v(t)$.

$$v(t) = V_0[1 + f(t)] \tag{7.112}$$

where

V_0 = the nominal recording velocity

$f(t)$ = the fractional flutter (an arbitrary function restricted only to having a zero mean value)

The distance, S(T), that the tape moves in a time interval T can be determined by integrating the velocity.

$$S(T) = \int_0^T v(t)\, dt$$

$$= V_0\left[T + \int_0^T f(t)\, dt\right] \tag{7.113}$$

Now the data can be expressed as a function of distance along the tape.

$$x(S) = x\left[V_0 T + V_0 \int_0^T f(t)\, dt\right] \tag{7.114}$$

If the tape is reproduced at a constant velocity V_1 (playback flutter is assumed equal to zero for simplicity), the output voltage is

$$x(t') = x^m(t) = x\left[\frac{V_0}{V_1}\, t + \frac{V_0}{V_1}\int_0^t f(u)\, du\right] \tag{7.115}$$

where

t' = machine time

$x^m(t)$ = the measured value on the output of the tape recorder at time t

If $V_1 = V_0$,

$$x^m(t) = X\left[t + \int_0^t f(u)\, du\right]$$

$$= X\left[t + g(t)\right] \tag{7.116}$$

where

$$g(t) = \int_0^t f(u)\, du \tag{7.117}$$

From Eq. (7.116) it can be seen that the flutter causes a time varying phase error in the reproduced version of the original data. That this error takes the form of frequency modulation of the data can be seen from Eq. (7.117). When the data is recorded on the tape by means of FM electronics, there are two additional flutter related errors that must be considered. These are a noise term from frequency modulation of the carrier frequency and an amplitude modulation of the data signal. These two errors will not be considered because (1) they can be eliminated by use of tape flutter compensation circuits and (2) Reference 49 has shown that the magnitude of these errors are insignificant for high quality instrumentation tape recorders even if flutter compensation is not used.

The development in Section 7.2 approximates the frequency modulated data signal shown in Eq. (7.116) by a two-term Taylor's series expansion. For this approximation to be valid, the third and higher order terms of the Taylor's series expansion must be negligible. The actual range over which this approximation is useful is dependent on the magnitude of the flutter and the frequency content of both the flutter and the data signal. In general, it can be said that the approximation will not be valid for very high data frequencies or for very low flutter frequencies. The two term approximation is shown below.

$$x[\, t + g(t)\,] \approx x(t) + g(t)\, \dot{x}(t) \tag{7.118}$$

For those cases where this approximation is valid, the autocorrelation function of the non-uniform tape velocity modulated data is shown in Eq. (7.55) to be

$$R_x^m(\tau) \approx R_x(\tau) - R_x''(\tau)\, R_g(\tau) \tag{7.119}$$

where

$R_x^m(\tau)$ = the autocorrelation function of the reproduced data signal

$R_x(\tau)$ = the autocorrelation function of the true data signal

$R_x''(\tau)$ = the second derivative, with respect to τ, of $R_x(\tau)$

$R_g(\tau)$ = the autocorrelation function of $g(t)$

By repeating the same development for a second data signal, $y(t)$, recorded on a second tape track one can show that

$$y^m(t) = y[t + h(t)] \tag{7.120}$$

$$R_y^m(\tau) \approx R_y(\tau) - R_y''(\tau)\, R_h(\tau) \tag{7.121}$$

$$R_{xy}^m(\tau) \approx R_{xy}(\tau) - R_{xy}''(\tau)\, R_{gh}(\tau) \tag{7.122}$$

where

$y^m(t)$ = the reproduced data signal

$h(t)$ = the dynamic phase error in the second tape track

$R_y^m(\tau)$ = the autocorrelation function of the reproduced version of $y(t)$

$R_y(\tau)$ = the autocorrelation function of $y(t)$

$R_y''(\tau)$ = the second derivative of $R_y(\tau)$ with respect to τ

$R_h(\tau)$ = the autocorrelation function of $h(t)$

$R_{xy}^m(\tau)$ = the cross-correlation function of the reproduced versions of $x(t)$ and $y(t)$

$R_{xy}(\tau)$ = the cross-correlation function of $x(t)$ and $y(t)$

$R_{xy}''(\tau)$ = the second derivative of $R_{xy}(\tau)$ with respect to τ

$R_{gh}(\tau)$ = the cross-correlation function of $g(t)$ and $h(t)$

The cross-correlation function between the dynamic phase errors can be further expanded as shown below:

$$g(t) - h(t) = \alpha(t) \tag{7.123}$$

where

$\alpha(t)$ = the differential dynamic phase error between the two tape tracks

$$R_{gh}(\tau) = R_h(\tau) + R_{\alpha h}(\tau) \tag{7.124}$$

where

$R_{\alpha h}(\tau)$ = the cross-correlation function between $\alpha(t)$ and $h(t)$

If the differential dynamic phase error is uncorrelated with the dynamic phase error $h(t)$, then the cross-correlation function $R_{gh}(\tau)$ becomes simply

$$R_{gh}(\tau) = R_h(\tau)$$

This fact is pointed out merely as a matter of interest. It is not known if $h(t)$ and $\alpha(t)$ are correlated, or not, in typical instrumentation recorders.

7.4.2 Analytical Examples

To examine the characteristics of the errors resulting in correlation measurements from recording the data on magnetic tape, several examples are worked assuming specific forms for both the flutter and the data. Errors in the ordinary spectral density and cross-spectral density measurements are not discussed since these can be established from the difference

between the Fourier transforms of the correlation functions of the actual data and the reproduced data. Also, since Eqs. (7.119) and (7.122) are of the same form, only the autocorrelation function computation errors will be explored.

Eq. (7.119) can be rewritten as

$$R_x^m(\tau) \approx R_x(\tau)\left[1 - \left\{\frac{R_x''(\tau)}{R_x(\tau)}\right\} R_g(\tau)\right] \qquad (7.125)$$

As long as the assumption in Eq. (7.118) is applicable, the error term in Eq. (7.125) can be split into two terms. The first term, $R_x''(\tau)/R_x(\tau)$, is related to the data only and the second term, $R_g(\tau)$, is related to the dynamic phase error only. The total error is of course the product of these two terms. The individual effects of these two terms will be examined separately.

First, consider the data related term $R_x''(\tau)/R_x(\tau)$. If the data is sinusoidal where $x(t) = A \sin \omega_1 t$

$$R_x(\tau) = \frac{A^2}{2} \cos \omega_1 \tau$$

$$R_x''(\tau) = -\frac{A^2}{2} \omega_1^2 \cos \omega_1 \tau$$

Hence,

$$\frac{R_x''(\tau)}{R_x(\tau)} = -\omega_1^2 \quad , \quad \text{for sinusoidal data} \qquad (7.126)$$

Thus the contribution of the data related term to the error in the correlation function increases with the square of the data frequency. This error is plotted in Figure 7-7.

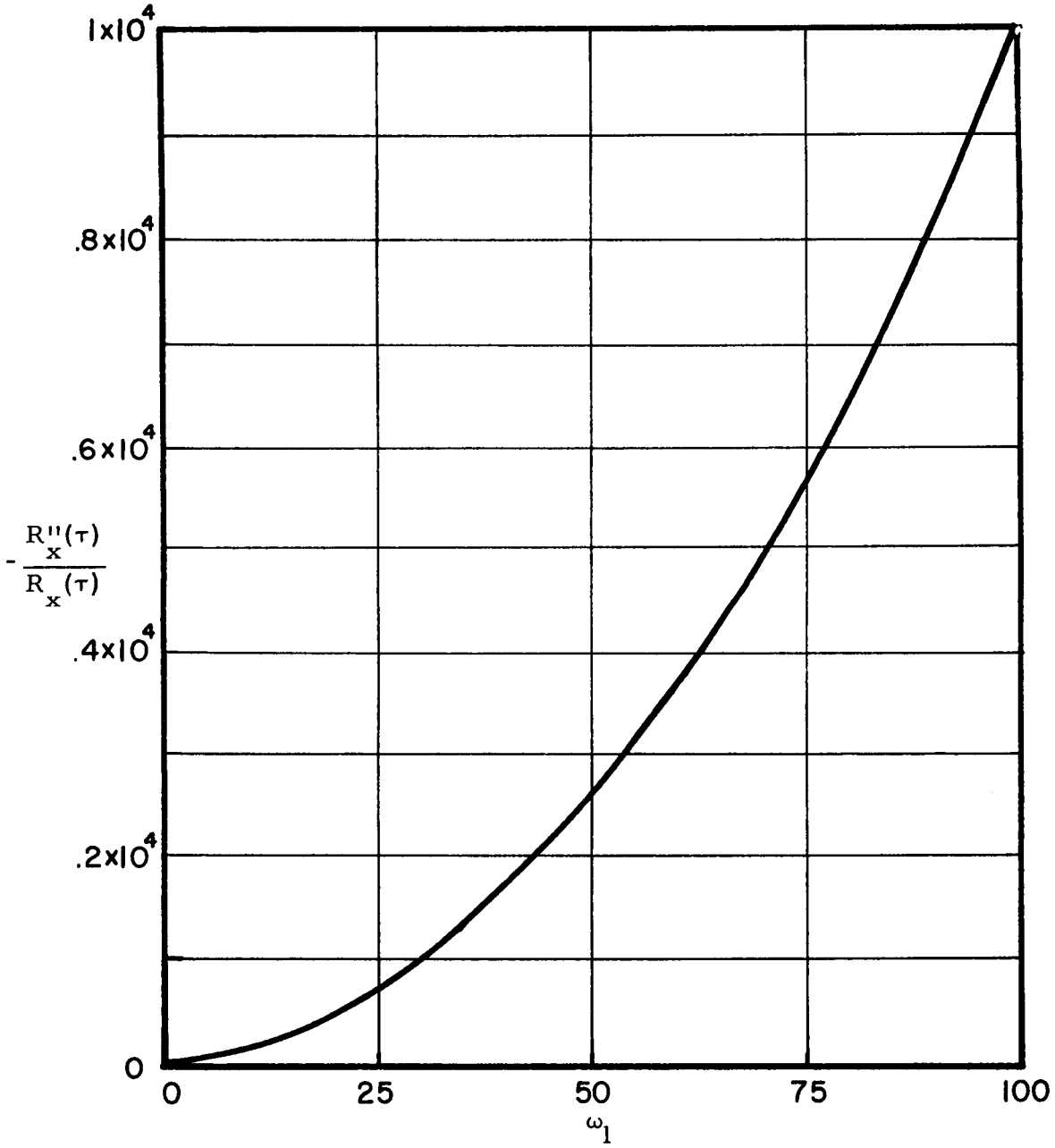


Figure 7-7. Data Related Phase Error Term for Sinusoidal Data

If the data consists of white Gaussian random noise passed through an ideal lowpass filter (see Figure 7-8), it can be shown that the data related error term is

$$\frac{R_x''(\tau)}{R_x(\tau)} = -\omega_1^2 - \frac{2\omega_1 \cot \omega_1 \tau}{\tau} + \frac{2}{\tau^2} \tag{7.127}$$

where ω_1 = the upper cutoff frequency of the lowpass filter. This error is plotted in Figure 7-9, for $\omega_1 = 100$.

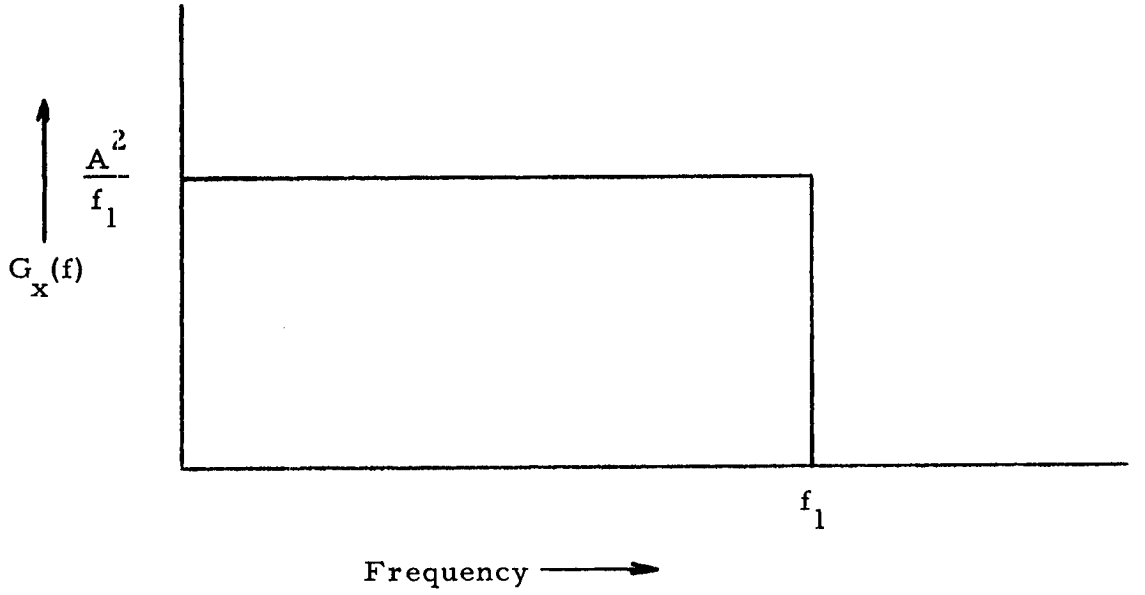


Figure 7-8. Spectral Density of Ideally Low Passed White Gaussian Noise

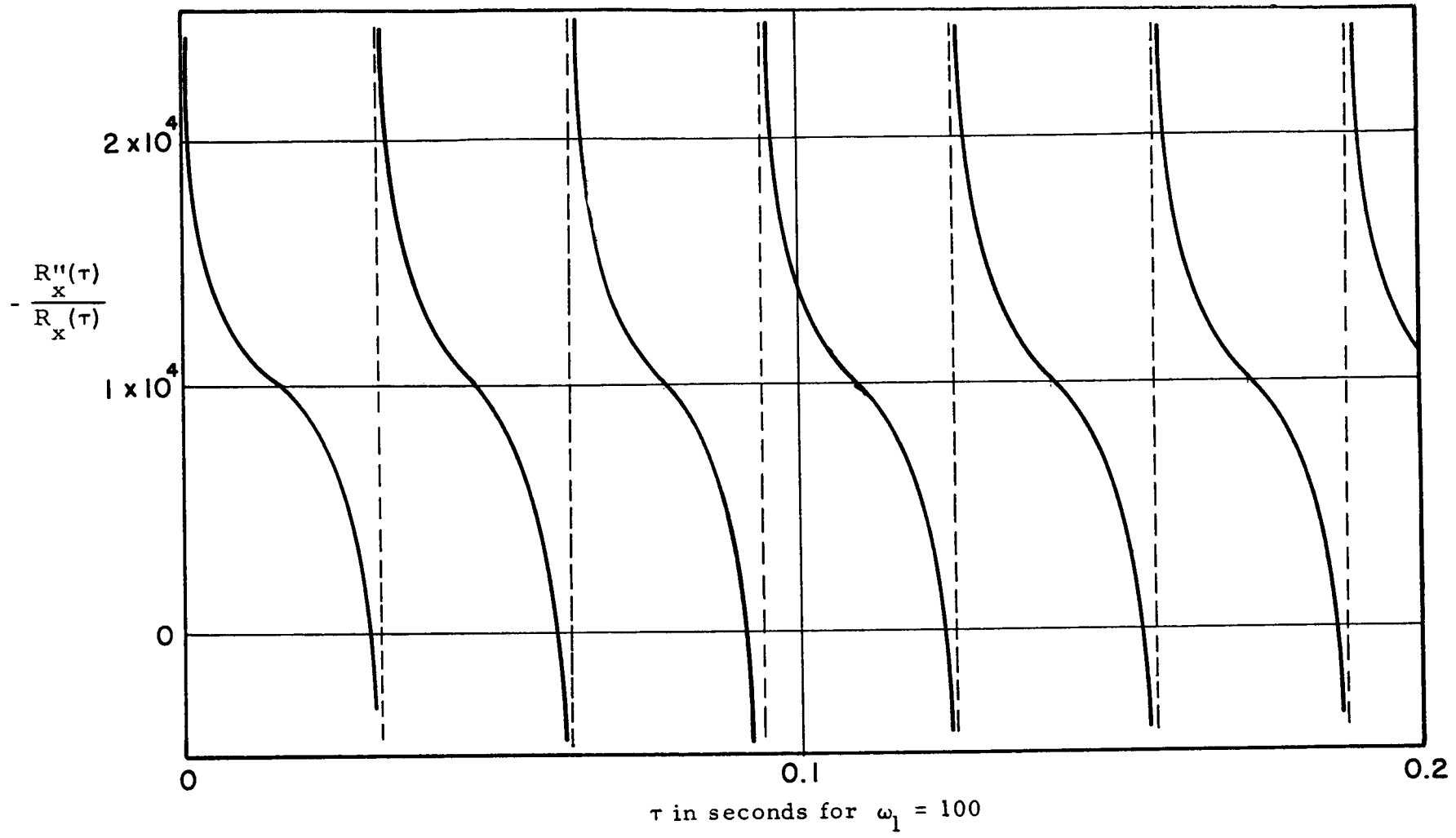


Figure 7-9. Data Related Phase Error for Data with an Ideal Lowpassed Spectral Density

For this case, the dependency on the square of the upper frequency has been modified by two additive terms. The first of these two terms ($2\omega_1 \cot \omega_1\tau/\tau$) tends to increase with the upper data frequency for a given time delay and tends to decrease with increasing time delay for a given cutoff frequency. The cotangent portion of this term reflects the particular normalization used. Since $R_x(\tau)$ passes through zero at intervals of $\omega_1\tau = \pi$, any error causes the normalized error to become infinite at intervals of $\omega_1\tau = \pi$. The second of the additive terms ($2/\tau^2$) decreases with the square of the increasing time delay.

If the spectral density of the data has the shape of the positive half of a Gaussian probability density function starting at f = 0 (see Figure 7-10), the data related error term can be shown to be

$$\frac{R_x''(\tau)}{R_x(\tau)} = -\left(\frac{\omega_1}{2.4}\right)^2 \left[1 - \left(\frac{\omega_1\tau}{2.4}\right)^2\right], \quad \text{for low passed Gaussian shaped spectral density} \tag{7.128}$$

where ω_1 = the half power frequency.

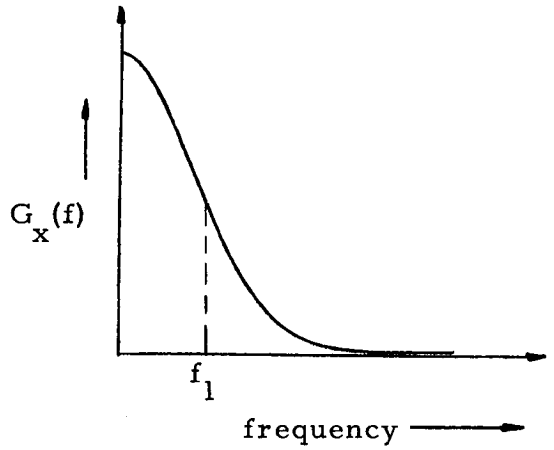


Figure 7-10. Gaussian Shaped Spectral Density

This error is shown in Figure 7-11, for $\omega_1 = 100$.

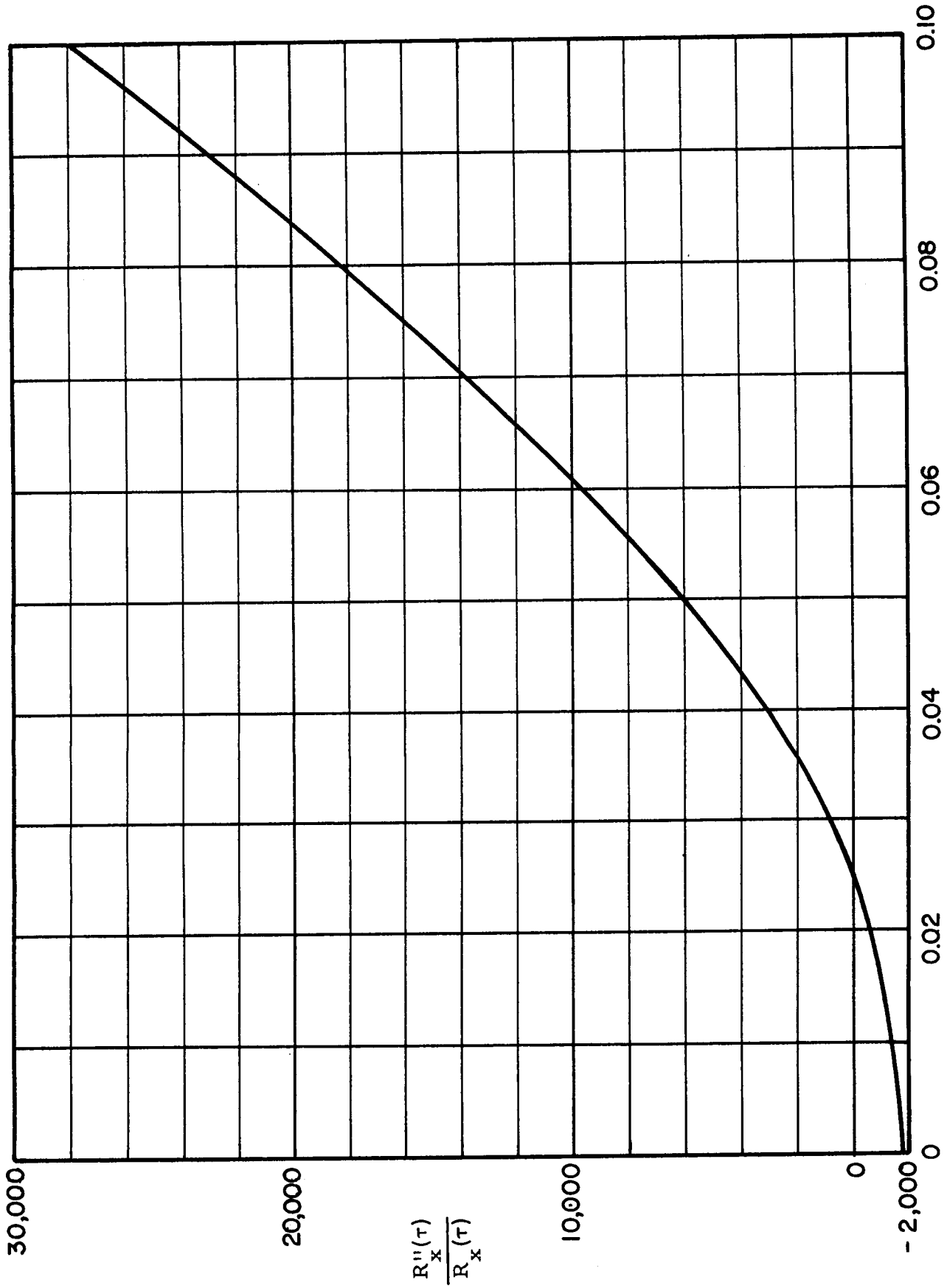


Figure 7-11. Data Related Phase Error Term for Data with Gaussian Shaped Spectral Density

For this case, the dependency on the square of the cutoff frequency has a scale factor $\left[\left(\frac{1}{2.4}\right)^2\right]$ and one additive term $\left[(\omega_1\tau/2.4)^2\right]$. This additive term increases as the square of the $\omega_1\tau$ product. However, since it is opposite in sign to the constant term, the total error will decrease to zero at $\omega_1\tau = 2.4$ and from that point on increase, with opposite sign.

Now consider the portion of the error term related to the tape speed non-uniformities, $R_h(\tau)$. This term is of major concern since it is descriptive of the error from the tape recorder. If this error can be made small enough, the total errors will be insignificant. Not much is known about the actual characteristics of the tape velocity non-uniformities. Since the tape is driven by rotary mechanisms, it is logical to assume that the flutter will contain some periodic components related to the rotational frequencies.

If it is assumed that the flutter is a single sinusoidal component, the phase error can be found from Eq. (7.117).

$$g(t) = \int_0^t f(u)\,du = \int_0^t B\cos\omega_2 u\,du$$

$$= \frac{B}{\omega_2}\sin\omega_2 t$$

$$R_g(\tau) = \frac{1}{2}\left(\frac{B}{\omega_2}\right)^2\cos\omega_2\tau \qquad (7.129)$$

This error is plotted in Figure 7-12. Equation (7.129) shows one very important fact. The higher the flutter frequency, ω_2, the lower will be the resulting error. In fact, the error term decreases with the square of the flutter frequency. This indicates that tape recorders having low mass,

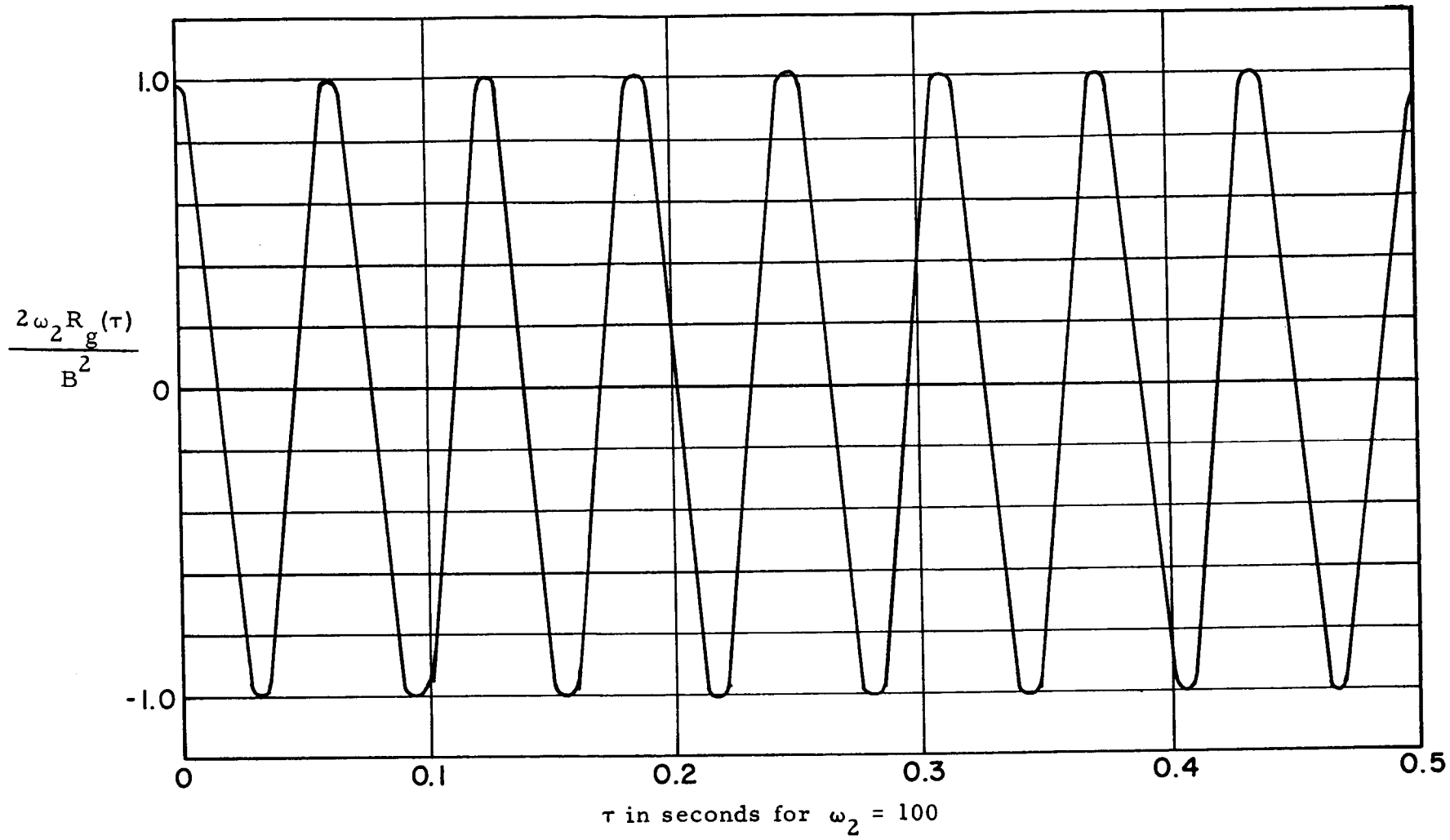


Figure 7-12. Flutter Related Phase Error for Sinusoidal Flutter

rapid response tape speed servo systems are preferable for correlation measurements because they filter out higher flutter frequencies than standard high mass tape drive systems.

If the flutter is random with a bandpass Gaussian white noise spectrum as shown in Figure 7-13,

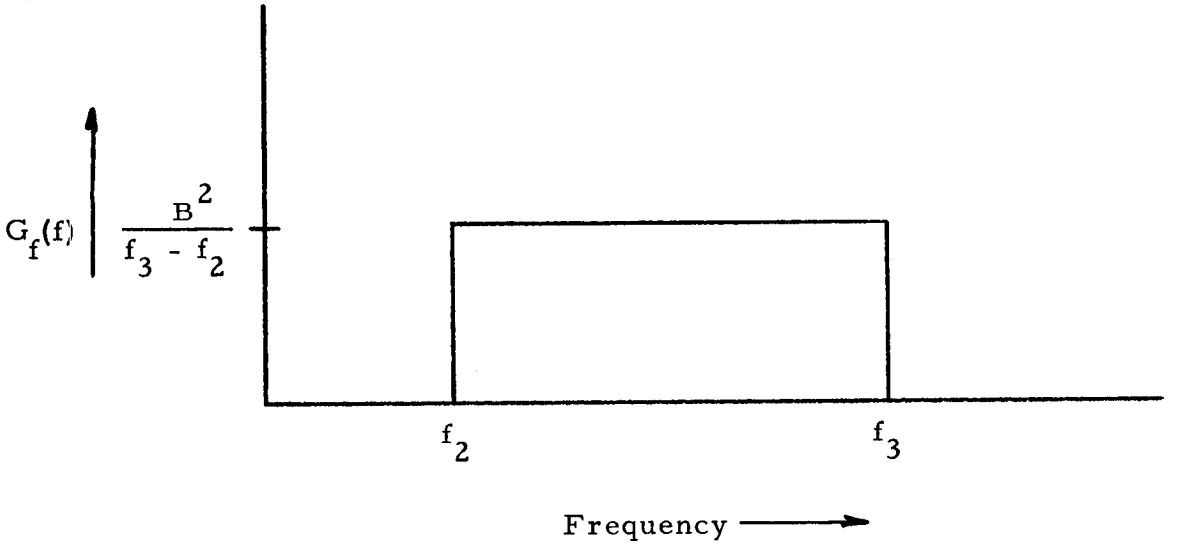


Figure 7-13. Ideal Bandpassed White Gaussian Flutter Spectrum

The dynamic phase error spectrum is as shown in Figure 7-14.

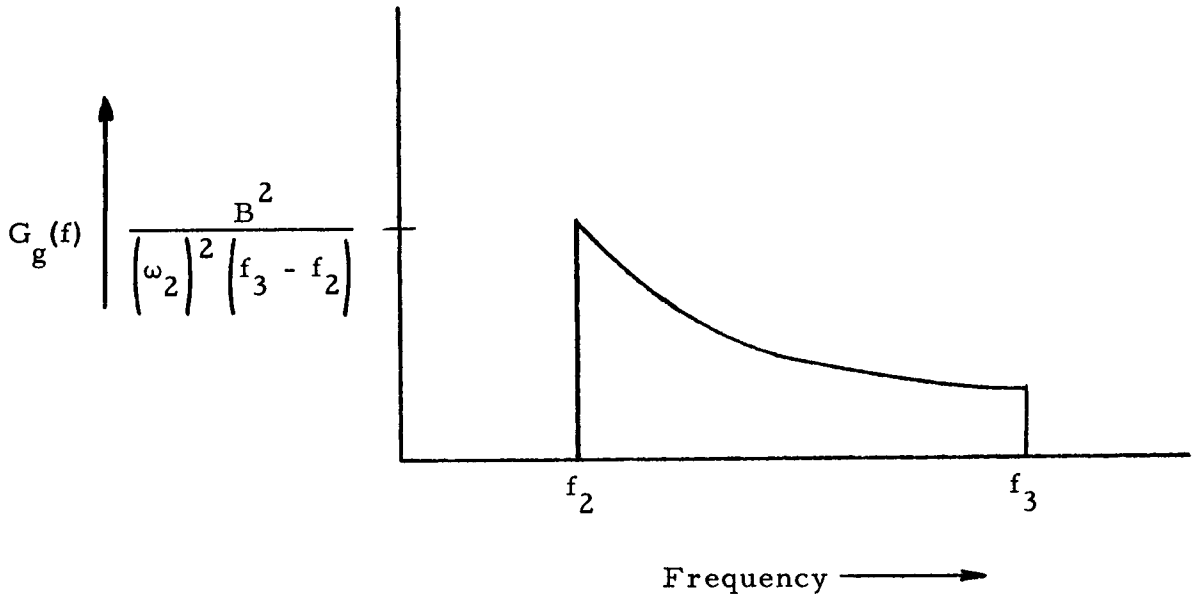


Figure 7-14. Dynamic Phase Error Spectrum Corresponding to Ideal Bandpassed Flutter Spectrum

Since $g(t) = \int f(t)\, dt$, $G_g(f) = \frac{1}{\omega^2} G_f(f)$

Hence $G_g(f) = \frac{1}{\omega^2}\left[\frac{B^2}{f_3 - f_2}\right]$; $f_2 \leq f \leq f_3$

$= 0$ elsewhere

The autocorrelation function of $g(t)$ can be found by taking the cosine transform of $G_g(f)$.

$$R_g(\tau) = \int_0^{\infty} G_g(f) \cos 2\pi f\tau \, df$$

$$= \left[\frac{B^2}{f_3 - f_2}\right] \int_{f_2}^{f_3} \left[\frac{\cos 2\pi f\tau}{(2\pi f)^2}\right] df$$

$$= \left[\left(\frac{1}{f_3 - f_2}\right)\left(\frac{B}{2\pi}\right)^2\right]\left[-\frac{\cos 2\pi f_3\tau}{f_3} + \frac{\cos 2\pi f_2\tau}{f_2} - 2\pi\tau\left\{\text{Si}(2\pi f_3\tau) - \text{Si}(2\pi f_2\tau)\right\}\right] \quad (7.130)$$

where $\text{Si}(\omega\tau)$ = the sine integral.

In a recorder suitable for recording high frequency acoustic data, one would expect f_3 to be 10 KHz or greater and f_2 to be less than 100 Hz. Under these conditions Eq. (7.130) can be approximated by

$$R_g(\tau) \approx \left[\left(\frac{1}{f_3}\right)\left(\frac{B}{2\pi}\right)^2\right]\left[\frac{\cos 2\pi f_2\tau}{f_2} - 2\pi\tau\left\{\pi/2 - \text{Si}(2\pi f_2\tau)\right\}\right] \quad (7.130)$$

This equation is plotted in Figure 7-15 as a function of τ for $f_2 = 10$ Hz. The higher f_2, the more rapidly as a function of τ, the oscillations decay. This also indicates that the flutter components should be filtered out to as high a frequency as is possible to perform accurate correlation analyses. To estimate the total error term one can graphically multiply the appropriate data related error term and flutter related term.

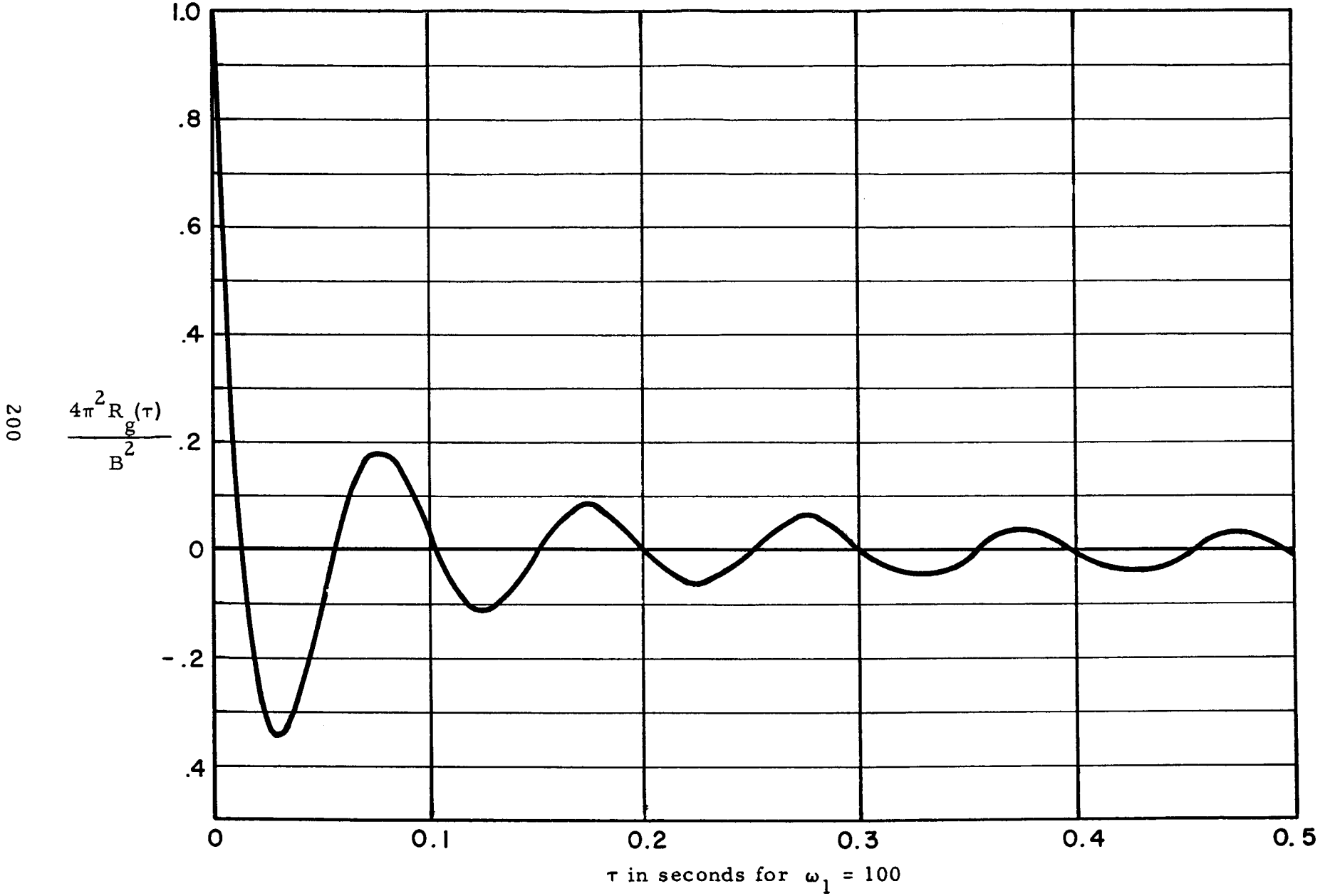


Figure 7-15. Flutter Related Phase Error Term for White Bandlimited Flutter

REFERENCES

1. Bendat, J. S. and A. G. Piersol, Measurement and Analysis of Random Data, John Wiley and Sons, Inc., New York, 1966.

2. Goff, Kenneth N., "An Analog Electronic Correlation for Acoustic Measurements," Journal of Acoustical Society of America, Vol. 27 No. 2, March 1955.

3. Gulton Industries Bulletin R2a, "Ortholog Laboratory Correlation Computer Model OCA-1."

4. Sangamo Electric Company Bulletin 3050, "Sangamo 480 Time Delay Capstan/Storage/Capstan Magnetic Tape Recorder Reproducer."

5. Conductron Corporation Bulletin, "Optical Products-Optical Correlator-Optical Analyzer-Optical Benches - Accessories."

6. Seiscor Division of Seismograph Service Corp. Bulletin, "Model VICH-100 Magnetic Correlator."

7. Kaiser, James F., and R. K. Angell, "New Techniques and Equipment for Correlation Computation," MIT Department of Electrical Engrg., Servomechanism Laboratory Technical Memorandum 7668-TM-2, December 1957.

8. Gardner, Murray F., and John L. Barnes, Transients in Linear Systems, John Wiley and Sons, Inc., New York 1942.

9. Bendat, J. S., Principles and Applications of Random Noise Theory, John Wiley and Sons, New York, 1958.

10. Parzen, E., "On Statistical Spectral Analysis," Proc. of the Symposia in Applied Mathematics, Vol. XVI, Stochastic Processes in Mathematical Physics and Engineering, American Mathematical Society, Providence, R. I., 1964.

11. Bendat, J. S., Enochson, L. D., and A. G. Piersol, "Analytical Study of Vibration Data Reduction Methods," NASA CR-55576 (N64-15529), National Aeronautics and Space Administration, Washington D. C., September 1963.

12. Simpson, S. M., Jr., "Time Series Techniques Applied to Underground Nuclear Detection and Further Digitized Seismic Data," MIT Scientific Report No. 2 Contract AF19(604)-7378, ARPA Order No. (80-61), December 8, 1961.

13. Weinreb, S., "A Digital Spectral Analysis Technique and its Application to Radio Astronomy," MIT Research Laboratory of Electronics, Technical Report 412, August 1963.

14. Bordner, G. W., Greaves, C. J. and W. W. Wierwille, "Research Studies of Random Process Theory and Physical Applications," NASA CR-61081, August 1964.

15. Schmid, L. P., "Efficient Autocorrelation," Communications of the ACM, Vol. 8, February 1965, p. 115.

16. Cooley, J. W. and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Mathematics of Computation, Vol. 19, No. 90, pp. 297-301, April 1965.

17. Arabadjis, C., "An Investigation of Linear Finite Optimum and Frequency Constrained Digital Filerts for Random and Deterministic Signals," Rev. 1, General Electric Defense System Dept. R62DSD56, January 1963.

18. Blackman, R. B., Linear Data-Smoothing and Prediction in Theory and Practice, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1965.

19. Blackman, R. B., and J. W. Tukey, The Measurement of Power Spectra, Dover Publications, New York, 1958.

20. Otnes, R. K., "Variable Bandwidth Spectral Computations," Measurement Analysis Corporation report MAC 404-04, March 1965.

21. Enochson, L. D., "An Automated Vibration Data Rejection System," Measurement Analysis Corporation Report MAC 404-03, November 1964.

22. Adage, Inc., Bulletin, "Ambilog 200."

23. General Applied Science Laboratories Bulletin, No. 1016, "Low Frequency Analysis," by M. Kaufman and P. Schaten, July 1964.

24. Piersol, A. G., "The Measurement and Interpretation of Ordinary Power Spectra for Vibration Problems," National Aeronautics and Space Administration, Washington, D. C., NASA CR-90(N64-30830) Sept. 1964.

25. Paul, R. J. A., "Hybrid Methods for Function Generation," The College of Aeronautics (Cranfield), Report No. 153, November 1961.

26. Handler, H., and Mangels, R. H., "A Delta-Sigma Modulation System for Time Delay and Analog Function Storage, "Electrical Engrg. Dept., University of Arizona, AD 434 539, January 1964.

27. Gatland, H. B., and R. J. A. Paul, "A Correlator Based on Delta Modulation Techniques," The College of Aeronautics (Cranfield) Report B and C, No. 3, August 1964.

28. Lee, W. H., "The Historical Development of Boeing Correlation Computers," Boeing Document No. D6-4285, December 1962.

29. Miyata, J., and T. Lentz, "MAGOP - A New Approach to High Density Digital Magnetic Recording," Large-Capacity Memory Techniques for Computing Systems, Edited by M. C. Yovits, Macmillan Company, New York, 1962.

30. Knoll, M., and B. Kazan, Storage Tubes and their Basic Principles, John Wiley and Sons, Inc., New York, 1952.

31. Bryan, J. S., and L. R. Rocht, "The Cathode Ray Tube as a Commutating Device in Large-Capacity, Random Access Stores," Large-Capacity Memory Techniques for Computing Systems, Edited by M. C. Yovits, New York, 1962.

32. Noble, D. L., "Magnetic Transducers and Amplifiers for Disk Recording," Large-Capacity Memory Techniques for Computing Systems, Edited by M. C. Yovits, Macmillan Company, New York.

33. Cerni, R. H., and L. E. Foster, Instrumentation for Engineering Measurement, John Wiley and Sons, New York, 1962, pp. 318-320.

34. Angelo, E. J., "An Electron-Beam Tube for Analog Multiplication, "Technical Report No. 249, October 1952, Research Laboratory of Electronics, MIT

35. Johnson, R. F., "A Time-Division Analog Multiplier for Correlation Measurements and Mixing at Frequencies up to 100 Kilocycles per Second,"NPL Aero Report 1030, August 1962, National Physical Laboratory, Aerodynamics Division, Teddington, Middlesex, England.

36. Levin, M. J., and J. F. Reintjes, "A Five Channel Electronic Analog Correlator," Proceedings of the National Electronics Conference, Vol. 8, 1952, pp. 647-652.

37. Page, R. M., and S. F. George, "Magnetic Drum Storage Applied to Surveillance Radar," Naval Research Laboratories Report 4878, January 1957.

38. Kelly, R. D., "Transducers for Sonic Fatigue Measurements," Air Force Flight Dynamics Laboratory Report, AFFDL-TR-64-171, February 1965.

39. Kelly, R. D., "Systems for the Collection and Analysis of Dynamic Data," Air Force Flight Dynamics Laboratory Report, AFFDL-TR-65-94, August 1965.

40. Choi, S. C., and L. D. Enochson, "Random Fatigue Test Sampling Requirements," Air Force Flight Dynamics Laboratory Report, AFFDL-TR-65-95, August 1965.

41. Middleton, D., An Introduction to Statistical Communication Theory, McGraw-Hill Book Company, Inc., New York, 1960.

42. Davies, G. L., Magnetic Tape Instrumentation, McGraw-Hill Book Company,Inc., New York, 1961.

43. Corcos, G. M., "Resolution of Pressure in Turbulence," Journal of the Acoustical Society of America, Vol. 35, No. 2, 1963.

44. Gilchrist, R. B. and W. A. Strawderman, "Experimental Hydrophone-Size Correction Factor for Boundary Layer Pressure Fluctuations," Journal of the Acoustical Society of America, Vol. 38, No. 2, 1965.

45. Willmarth, W. W. and F. W. Roos, "Resolution and Structure of the Wall Pressure Field Beneath a Turbulent Boundary Layer," Journal of Fluid Mechanics, Vol. 22, 81-94, 1965.

46. Chandiramani, K. L., "Interpretation of Wall Pressure Measurements Under a Turbulent Boundary Layer," BBN Report 1310, August 1965.

47. Schulze, G. H., "Tape Recording Errors," ISA Journal, Vol. 11, No. 5, May 1964.

48. Wise, J. H., "The Tape Recorder as a Source of Phase Error," Journal of Institute of Environmental Sciences, June 1965.

49. Ratz, A. G., "The Effect of Tape Transport Flutter on Spectrum and Correlation Analysis," IEEE Transactions on Space Electronics and Telemetry, Vol. SET-10, No. 4, December 1964.

50. Chao, S. C., "The Effect of Flutter on a Recorded Sinewave," Proceedings of the IEEE Correspondence, July 1965.